



PhD-FDEF-2011-05

Fakultät für Rechts-, Wirtschaft- und Finanzwissenschaften

## DISSERTATION

verteidigt am 27/06/2011 in Luxemburg

zur Erlangung des Titels

## DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN MATHÉMATIQUES

von

**Stefan Giebel**

geboren am 7.Juni 1976 in Wolfhagen

## ZUR ANWENDUNG DER STATISTISCHEN FORMANALYSE

### Prüfungskommission

Dr., Betreuer der Doktorarbeit,  
Jang Schiltz  
*Professor, Universität Luxemburg*

Dr., Vorsitzender  
Raymond Bisdorf  
*Professor, Universität Luxemburg*

Dr., Stellvertretender Vorsitzender  
Franz-Bernd Frechen  
*Professor, Universität Kassel*

Dr.,  
Jens-Peter Schenk  
*PD, Universitätsklinikum Heidelberg*

Dr.  
Martin Rainer  
*Professor, METU Ankara*



# Zur Anwendung der statistischen Formanalyse

Stefan Markus Giebel

29. Juli 2011

---

# Inhaltsverzeichnis

<b>1</b>	<b>Abstract</b>	<b>5</b>
<b>2</b>	<b>Einleitung</b>	<b>7</b>
<b>3</b>	<b>Die statistische Formanalyse</b>	<b>13</b>
3.1	Der Gedanke der "Form" . . . . .	13
3.2	Eine kurze Historie der Formanalyse . . . . .	17
3.3	Die Grundbegriffe der statistischen Formanalyse . . . . .	20
3.4	Die mittlere Form . . . . .	30
3.5	Berechnung der mittleren Form . . . . .	36
3.6	Varianz . . . . .	38
3.7	Mittlere Varianz . . . . .	39
3.8	Alternative Berechnung einer mittleren Form mittels neuronaler Netze . . . . .	40
<b>4</b>	<b>Differenzierung und Relevanz der Landmarken</b>	<b>43</b>
4.1	Wiederherstellung von Formen . . . . .	43
4.2	Test auf Differenzierbarkeit . . . . .	44
4.2.1	Test der mittleren Form . . . . .	44
4.2.2	Test der Varianz . . . . .	47
4.2.3	Test der mittleren Form und der Varianz . . . . .	48
4.2.4	Entscheidungsbäume . . . . .	49
4.3	Landmarkentest . . . . .	52
4.3.1	Test auf Einfluss der Landmarken auf die Form . . . . .	52
4.3.2	Selektion einer Auswahl an Landmarken . . . . .	54
4.3.3	Vorzeichenrangtest . . . . .	56
<b>5</b>	<b>Funktionale Formanalyse</b>	<b>63</b>
5.1	Mittlere zu erwartende Funktion . . . . .	63
5.2	Dynamische Formanalyse . . . . .	67
<b>6</b>	<b>Neuronale Netze</b>	<b>71</b>
6.1	Elemente Neuronaler Netze . . . . .	71
6.2	Mehrschichtiges neuronales Netz . . . . .	73

---

6.3	Neuronale Netze in Kombination mit stochastischen Prozessen . . . . .	74
6.4	Neuronale Netze in Kombination mit der Formanalyse . . . . .	76
6.4.1	Neuronale Netze zur Differenzierung der Gruppen . . . . .	77
<b>7</b>	<b>Klassifikation</b>	<b>81</b>
7.1	Einführung . . . . .	81
7.2	Logistische Regression . . . . .	82
7.3	Diskriminanzanalyse . . . . .	85
7.4	Chaotisches Verhalten der verwendeten Verfahren . . . . .	89
7.5	P-KFA . . . . .	91
<b>8</b>	<b>Anwendungen der Formanalyse</b>	<b>93</b>
8.1	Nierentumore im Kindesalter . . . . .	94
8.2	Räumliche Analyse des menschlichen Verhaltens . . . . .	97
8.3	Elektronische Nasen . . . . .	98
8.3.1	Beschreibung der Geruchsqualität durch Unterscheidung von Standorten . . . . .	99
8.3.2	Detektion von Drogen und Sprengstoff über die Geruchsqualität . . . . .	99
8.3.3	Bestimmung der Geruchsquantität bzw. Geruchsbelastung	101
<b>9</b>	<b>Auswertung der Daten</b>	<b>103</b>
9.1	Nierentumore bei Kleinkindern . . . . .	103
9.1.1	Erfassung der Daten . . . . .	103
9.1.2	Test auf Unterscheidbarkeit der Gruppen . . . . .	107
9.1.3	Test der Landmarken zur Unterscheidung der Gruppen . . . . .	113
9.1.4	Ergebnisse der Klassifikation . . . . .	114
9.2	Räumliche Analyse bei Serienmördern . . . . .	118
9.2.1	Erfassung der Daten . . . . .	118
9.2.2	Auswertung . . . . .	119
9.2.3	Test auf Unterscheidbarkeit der Gruppen . . . . .	119
9.3	Sensordaten elektronischer Nasen . . . . .	122
9.3.1	Messungen . . . . .	122
9.3.2	Test auf Unterscheidbarkeit der Gruppen . . . . .	122
<b>10</b>	<b>Fazit</b>	<b>125</b>
10.1	Formanalyse . . . . .	125
10.2	Datenerfassung . . . . .	125
10.3	Ungewissheit der Messung . . . . .	126
10.4	Mathematische Verfahren . . . . .	126
10.4.1	Selektion der Modelle . . . . .	126
10.4.2	Vorhersage . . . . .	127
10.4.3	Ausblick . . . . .	128

# Kapitel 1

## Abstract

In the following work we propose an introduction to Statistical Shape Analysis and adapt the procedures to three and four dimensional objects. Especially, in the case of four dimensional objects we show one possible approach of Dynamical Shape Analysis. In Shape Analysis we use parameters of the distribution as the mean shape, a representative shape of a group, to test the possibility of discriminating different kind of groups of objects. First-time we are using in the concept of Shape Analysis the variance. Furthermore we combine the Neural Networks with the Shape Analysis and also the tests used for discriminating groups with Answer Tree. To classify the objects well known procedures as Logistic Regression, Neural Networks, Discriminant Analysis and Answer Tree are used. For the application more than one hundred renal tumors in childhood, also more than one hundred electronic nose sensor data for different kind of odor quality and concentration as well as the behaviour of around fifty sexual serial murderer are measured and used. In all the disciplines a new approach in measurement of data is shown. The work shows furthermore, that all three fields Shape Analysis is applicable. The results of Shape Analysis and the classification afford the discrimination of different kind of objects. Shape Analysis could be one answer for classifying image data in the light of increasing image data.

---



# Kapitel 2

## Einleitung

In einer Vielzahl von Disziplinen ist es von großem praktischen Interesse, die Form von Objekten nachvollziehbar zu beschreiben und miteinander zu vergleichen. Allgemein kann die Form eines Objekts, einer Datenmenge oder eines Bildes als die Information definiert werden, die invariant gegenüber Translationen, Rotationen und Skalierung ist. Das Feld der Formanalyse beinhaltet daher die Methoden zur Auswertung von Formen, bei denen die Lage, die Größe und die Rotation eliminiert werden kann.

Die zwei- oder mehrdimensionalen Objekte werden durch entsprechende Punkte, so genannte Landmarken beschrieben. Die Landmarken ergeben sich entweder rein explorativ oder aufgrund der Erkenntnisse der Fachdisziplin.

Dieser Ansatz einer durch Landmarken beschriebenen Form führt uns zunächst schon allein optisch zu der Aussage, ob dadurch eine Differenzierung der Objekte überhaupt möglich scheint. Bei entsprechender Differenzierbarkeit gelangt man durch die Anwendung des Verfahrens zu einer Zuordnung der gegebenen Form zu einer vorgegebenen Gruppe von Objekten und damit zu einer objektiven Methode für die Klassifikation von Objekten, die gerade in der heutigen Zeit eine Alternative zu den zunehmend intuitiv erscheinenden Entscheidungsprozessen bei der Bewertung von Bildmaterial sein könnten<sup>1</sup>.

Für die Klassifikation bei vorgegebener Zuordnung kommt keine Clusteranalyse in Frage, da die Clusteranalyse lediglich eine Gruppenzuordnung anhand der Distanz oder Ähnlichkeit der Objekte untereinander vornimmt [57].

Sollte nur die Information der Form in das Verfahren eingehen, so wäre davon auszugehen, dass die Formen innerhalb ein- und derselben Gruppe nur gering voneinander abweichen, während die eine Gruppe bildenden Formen im Vergleich zu einer anderen Gruppe von Formen erheblich von dieser abweichen [32]. Sobald darüber hinaus bei den Formen den jeweils einzelnen Landmarken eine unterschiedliche Bedeutung bei der Zuordnung zu einer Gruppe beizumessen ist, ist die Clusteranalyse, die eine solche Bewertung und Gewichtung der Landmarken nicht anhand einer vorgegebenen Zielgröße vornimmt, nicht mehr

---

<sup>1</sup> Siehe [26] S.126

---

verwendbar [38] oder nur in Kombination mit einer entsprechenden Gewichtung der jeweiligen Landmarken.

Die statistische Formanalyse beschäftigt sich methodisch mit der Analyse von Formen unter Beachtung der Zufallsprozesse. Die statistische Formanalyse ist ein mathematisches Verfahren, um zwei- oder mehrdimensionale Daten unter Beachtung der Größe und der Lage des zu beschreibenden Objektes auszuwerten. So können Objekte unterschiedlicher Größe und Lage miteinander verglichen werden. Um die Form eines Objektes zu bestimmen, werden die Objekte standardisiert, zentriert usw. in einem dafür geeigneten metrischen Raum. Dieser Ansatz ist zu unterscheiden von der Wiederherstellung von Formen innerhalb von Bildern [10] und ist durch die Vorverarbeitung der Daten eine Erweiterung gegenüber einer reinen Berechnung von Distanzen zwischen Objekten [17].

In dieser Arbeit werden die für die Anwendung notwendigen theoretischen Ansätze der statistischen Formanalyse vorgestellt und praktische Anwendungen in verschiedenen Gebieten durchgeführt. Die bestehende Theorie wird hinsichtlich der Metrik, der Berechnung der mittleren Form erweitert und die mehrdimensionale mittlere Abweichung eingeführt. Außerdem werden vereinfachte Beweise für einige bestehende Resultate aufgezeigt.

Die verwendete euklidische Metrik setzt den zu betrachtenden Formveränderungen Grenzen. In Folge der verwendeten Metrik werden Unterschiede beachtet oder vernachlässigt. Daher ist gerade bei der Differenzierung von Objektgruppen stets eine für die Anwendung geeignete Metrik zu wählen. Man muss sich auch über die sich auf alle Landmarken im gleichem Umfang auswirkende Transformation im klaren sein und eventuelle Besonderheiten beachten (zum Beispiel Tumorgewebe im Bereich von weichem Gewebe wie Organe und in der Nähe der Körperbegrenzungen).

Die mittlere Form ist der Erwartungswert nach Fréchet [62]. Sie ergibt sich als die Form mit dem für die verwendete Metrik geringsten Abstand zu allen Objekten.

Der Erwartungswert der mittleren Form einer Gruppe von Objekten kann zur Differenzierung von Objektgruppen dienen. Außerdem wird die Formanalyse in Kombination mit Neuronalen Netzen durchgeführt, die im Rahmen des zu untersuchenden Wertebereichs ein Minimum für ein zu definierendes Maß wie die Anzahl an Fehlzuordnungen bestimmen können.

Darüber hinaus werden Ansätze vorgestellt, die Formen der Objekte mittels einer Funktion zu interpolieren und die Interpolationsfunktionen von Objekten für die Bestimmung einer Funktion verwendet, die die mittlere Form beschreibt. Dieser Ansatz macht nur dann einen Sinn, wenn eine entsprechende Vorkenntnis der Art der Verbindungen zwischen den Punkten existiert und von einer entsprechenden Bewertbarkeit und Deutbarkeit der Verbindungen zwischen den Punkten gesprochen werden kann.

---

Die mehrdimensionale mittlere Abweichung stellt analog zur mittleren Form die Beachtung von Rotation und Translation im mehrdimensionalen Raum sicher. Sie veranschaulicht die Abweichungen entlang der zu betrachtenden Dimensionen. Allein aus messtechnischen Gründen ist nicht unbedingt davon auszugehen, dass in allen Dimensionen und bei allen Landmarken eine gleiche Abweichung zu erwarten ist.

Aufbauend auf der Formanalyse werden Tests entwickelt, um die Unterschiede zwischen Gruppen von Objekten zu prüfen. Eine mittlere Form und mittlere Abweichung einer Gruppe stellt noch keine Differenzierung von anderen Gruppen sicher. Die Tests benutzen dabei den Abstand von der mittleren Form, die mehrdimensionale mittlere Abweichung, sowie den Abstand und die Varianz des Abstandes von der mittleren Form als Unterscheidungsmerkmale zwischen den Objektgruppen.

Um zwei Objektgruppen voneinander unterscheiden zu können, kann man analog zu dem Mann-Whitney-U-Test [90] die Abweichung von der mittleren Form, die von einer der beiden Objektgruppen berechnet wird, verwenden. Es wird davon ausgegangen, dass Objekte, die zur Berechnung der mittleren Form beigetragen haben, eine geringere Distanz zur mittleren Form haben als andere Objekte. Der Test von Ziezold [90] verlangt die Berechnung der mittleren Form der zu untersuchenden und der komplementären Gruppe, da bei ähnlichem Mittelwert der beiden Gruppen das Testergebnis, welches zu einem Verwerfen der Annahme der Gleichheit der Verteilungen berechtigt, eine Folge der Varianz sein kann, denn es wird die jeweilige "mittlere Form" einer Gruppe im Verhältnis zu einer anderen Gruppe geprüft. Ist andersherum die Varianz einer der beiden Gruppen sehr groß, so können sich die Objekte, die zu der Gruppe mit der geringeren Varianz gehören, in die andere Gruppe einfügen und allein dadurch zu einer schlechteren Differenzierbarkeit führen.

Eine Berechnung der mittleren Form einer Objektgruppe ist bezüglich der Anwendung auch nur dann sinnvoll, wenn von einer homogenen Gruppe ausgegangen werden kann. In vielen Anwenderfällen ist lediglich eine oder mehrere Gruppen von Objekten von allen anderen zu unterscheiden. Die verbliebene restliche Gruppe ist dann allein aus Anwendersicht zu heterogen, um von einer mittleren Form zu sprechen oder von einer solchen ausgehen zu können.

Die mehrdimensionale mittlere Abweichung dient zur Prüfung der Varianzunterschiede in den beiden Gruppen, ohne dass man von beiden Objektgruppen die mittlere Form zu berechnen braucht. Unter Verwendung der Ansätze des F-Tests wird geprüft, ob die Annahme der Varianzgleichheit verworfen werden kann.

Ein weiterer Ansatz zur Prüfung der mittleren Form ist die Einbeziehung der Varianz in den Test in Form der Standardabweichung  $\mu \pm \sigma$ . Die Tschebys-

---

cheffsche Ungleichung<sup>2</sup> ermöglicht verteilungsunabhängig die Einschätzung der eingegangenen Abweichungen von  $\mu$  anhand von  $P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$ , für alle reellen  $k > 0$ . Durch Veränderung von  $k$  kann die für die Differenzierung der Objektgruppen optimale zulässige Abweichung bestimmt werden.

Ein zu der Berechnung beider mittlerer Formen weiterer Nachteil des Tests von Ziezold [90] ist seine Anwendung lediglich im Vergleich zweier Gruppen zueinander. Ein Nachteil der durch ein schrittweises bivariates Vorgehen analog beispielsweise zum  $\chi^2$ -Test bei den Entscheidungsbäumen ausgeglichen werden kann und durch die Kombination des Tests mit dem schrittweise bivariaten Vorgehen eine weitergehende Differenzierung der Gruppen ermöglicht.

Für den Anwender ergibt sich des Weiteren die Fragestellung nach der Relevanz der Landmarken. Wir stellen drei verschiedene Landmarken-Tests vor, ein Test zur Messung eines eventuellen Einflusses der Landmarken auf die mittlere Form, ein Test zur Unterscheidbarkeit der Objektgruppen anhand der Landmarken und eine explorative Selektion und Prüfung von Landmarken.

Der Einfluss der Landmarken auf die Form wird durch den Test von Ziezold(2003) [91] untersucht. Ziezold selbst spricht dabei von der Prüfung auf Unabhängigkeit der Landmarken. Um Verwechslungen zwischen der stochastischen Unabhängigkeit und der von Ziezold verwendeten Begrifflichkeit zu vermeiden, sprechen wir lediglich von dem Einfluss der Landmarken.

Unabhängig von der Gruppengröße bietet der Wilcoxon-Test die Möglichkeit die Distanz der Landmarken von Objekten zur jeweils mittleren Form auf Gruppenunterschiede hin zu prüfen. Durch das Vorgehen kann man die Landmarken erkennen, die zu einer hohen Distanz zur mittleren Form beitragen.

Bei der explorativen Selektion von Landmarken wird eine bestimmte Anzahl von Landmarken ausgewählt und auf ihre Möglichkeit zur Differenzierbarkeit der Objektgruppen geprüft. Für die endgültige Prüfung der sich als relevant erwiesenen Landmarken werden noch unbekannte Daten verwendet oder der entsprechende Test zur Unterscheidung der Gruppen verwendet, wenn die Heterogenität einer der beiden Gruppen einem Datensplit entgegensteht.

Hinsichtlich der dynamischen Formanalyse wird im Weiteren lediglich ein Verfahren vorgestellt, welches die Veränderung von dreidimensionalen Körpern in der Zeit untersucht und vergleicht. Zusätzlich zur Formanalyse ist im Fall einer vierdimensionalen Betrachtung eine Funktion  $g$  zu ermitteln, die die Veränderung beschreibt. Um Gruppen zu unterscheiden, ist dann die Differenz der Transformationen über die Funktion  $g$  für das jeweilige Objekt zu betrachten und mit der Differenz der Transformationen über die Funktion  $\hat{g}$  der mittleren Form zu vergleichen. Wir können leider nur die Theorie dieses Verfahrens vorstellen, da bis zum Abschluss der Untersuchung keine für eine Aussage notwendige Datenmenge zusammengestellt werden konnte.

Für die endgültige Klassifikation der Objekte wird die Diskriminanzanalyse nach Fisher, die Logistische Regression, die Neuronale Netze und die Prädiktions-Konfigurationsfrequenzanalyse verwendet. Die Diskriminanzanaly-

---

<sup>2</sup> S. 116, [20]

---

se nach Fisher trennt die Gruppen linear voneinander, die Logistische Regression bestimmt eine Wahrscheinlichkeit für die Gruppenzugehörigkeit bei den entsprechenden Ausprägungen des Patienten und trennt nichtlinear. Die verwendeten zweischichtigen Neuronalen Netzen sind lediglich eine Erweiterung der Logistischen Regression unter Beachtung von Wechselwirkungen der Eingangsvariablen, insofern wie in vielen Disziplinen die Wechselwirkungen nicht bekannt sind. Im Folgenden wird zusätzlich auf das chaotische Verhalten in Neuronalen Netzen eingegangen und auf den Einfluss des chaotischen Verhaltens auf die Interpretation der Ergebnisse.

Die Prädiktions-Konfigurationsfrequenzanalyse setzt eine Diskretisierung der Eingangsvariablen voraus und bestimmt anhand der aus mehreren Merkmalen gebildete Profile die wahrscheinliche Zuordnung eines Profils zu einer Kategorie.

Unsere Arbeit ist folgendermaßen gegliedert: Zunächst wird der Grundgedanke der Form eines Objekts behandelt und der erste Einstieg in die statistische Formanalyse gewährt, dann wird den Tests zur Prüfung der Unterschiede der Objektgruppen nachgegangen, Klassifikationsverfahren vorgestellt und schließlich nach Beschreibung der Anwenderprobleme die statistischen Tests und die Klassifikationsverfahren auf die zu lösenden Probleme angewandt. Am Ende wird kritisch den Ergebnissen und Aussagen unserer Untersuchung nachgegangen.

---

## Kapitel 3

# Die statistische Formanalyse

### 3.1 Der Gedanke der "Form"

Bereits bei dem griechischen Mathematiker und Philosophen Plato [66] findet man in der Beschreibung des Ideals die Idee der "Form"; einer Form, die repräsentativ für eine Gruppe von Objekten steht, oder einer Schablone, mit der der Mensch immer wieder von neuem Objekte einordnen kann. Diese für die Gruppe von Objekten allgemeingültige Form ist zur höheren über die irdischen Dinge hinausgehenden Erkenntnisfindung zu suchen und zu beschreiben. Die von Plato angestrebte Erkenntnis hat noch einen religiösen Charakter im Sinne der Suche nach dem Ursprung alles Seienden. Wer sich nach ihm dem Ursprung der Erkenntnisfähigkeit nähert, nähert sich dem Göttlichen. In Platos Theorie der Formen wird behauptet, dass vielmehr die nicht-materiellen abstrakten Formen und nicht die veränderliche materielle Welt die Wirklichkeit beschreiben. Die Veränderungen sind bei ihm nur unterschiedliche Erscheinungen ein- und derselben Form. Alle von uns wahrgenommenen Dinge entsprechen nach ihm solchen Urbildern. Plato vertritt die Auffassung, dass diese Formen bzw. Urbilder die einzig wahren Objekte wären. Sie ermöglichen uns erst die Einordnung von wahrgenommenen Dingen in Kategorien und sie existieren nach Plato unabhängig von unserer Wahrnehmung. Ein Pferd wird somit als "Pferd" wahrgenommen, weil es dem Urbild eines "Pferdes" entspricht. Mittels der Urbilder bzw. Formen versucht Plato seine Lösung für das Problem des Universellen zu formulieren. Philosophisch gesehen ist mit H. Reichenbach [63] die Existenz der Dinge nur eine Frage der Perspektive und mit M. Heidegger [52] sind die Dinge erst in der "Jemeinigkeit" und in ihrer "Zuhandenheit" erfahrbar. Reichenbach stellt sich ausgehend von der Relativitätstheorie und der Quantenphysik nämlich die Frage, ob die Dinge, die wir momentan nicht wahrnehmen, überhaupt vorhanden sind. Er geht der Frage nach, ob der Baum in dem Moment noch existiert, wenn wir ihn nicht gerade wahrnehmen. Die Berechtigung der Frage ergibt sich aus

---

der physikalischen Erkenntnis der Abhängigkeit der Wahrnehmung des Lichts vom Beobachter als Teilchen oder als Welle.

Heidegger verweist auf die individuelle Wahrnehmung und die Unmöglichkeit, Dinge als solches wahrzunehmen, die nicht mit uns als in der Welt Handelnde im Zusammenhang stehen. Nur in unserer je eigenen Welt werden Dinge für uns erfahrbar. Objektivität an sich stützt sich auf Intersubjektivität, auf eine zwischen mindestens zwei Subjekten vergleichbare Erfahrung und Zuhandenheit der Dinge.

Hinsichtlich der mathematischen Herangehensweise ist unabhängig von dem objektiven Vorhandensein der Urbilder eine "Form", die zur Beschreibung der Wirklichkeit taugt und zur Einordnung weiterer Objekte dient, stets erstrebenswert, womit wir im Weiteren dem Pragmatismus von C.S. Peirce [65] in der Anwendung folgen. Es geht nicht darum, die Wahrheit hinter den Dingen zu finden, sondern darum ein Vorgehen zu finden, das sich als nützlich und handhabbar im Umgang mit der Welt erweist und eine Vergleichbarkeit im Umgang mit den uns erfahrbaren Dingen sicher stellt. Der Focus des Interesses wissenschaftlichen Denken und Handelns beeinflusst nicht nur die Gesellschaft und das Leben des Menschen, sondern die Gesellschaft und das Leben des Menschen bedingen den Focus allen Denkens.

Neben dem Tastsinn, ist das Auge das wichtigste Sinnesorgan des Menschen, wenn es um die Erkennung der Form geht. Das Auge ist in der Lage, die elektromagnetische Strahlung, im Wellenlängenbereich des Sonnenspektrums wahrzunehmen [18]. Das menschliche Auge in seiner Perfektion nimmt das gesamte Farbspektrum wahr und stellte für Kritiker der Darwinschen Evolutionslehre stets ein Beweis dafür da, dass es nicht durch zufällige Mutationen entstanden sein könnte. Mittlerweile gilt es als bewiesen, dass diese besonderen Zellen im Auge aus lichtempfindlichen Zellen des Gehirns entstanden sind [4].

Die Wahrnehmung an sich entsteht aus der Verarbeitung der Informationen von Mustern und Formen. Dabei werden die wahrgenommenen Punkte im Zusammenhang gesehen. Der Zusammenhang basiert auf den Erfahrungen des Menschen mit seiner Umwelt. Dies kann auch zu Fehleinschätzung der wahrgenommenen Informationen führen, wie dies in vielen optischen Täuschungen nachzuweisen ist. So sehen wir im folgenden Bild (3.1) graue Punkte zwischen den schwarzen Flächen, obwohl dort keine sind.

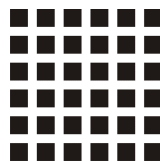


Abbildung 3.1: Beispiel einer optischen Täuschung



---

Mustererkennung ist die Fähigkeit, eine Menge von Daten hinsichtlich ihres Informationsgehalts durch Regeln, Wiederholungen oder Ähnlichkeiten verarbeiten zu können. Diese Leistung kennzeichnet im Sinne der Biologie höher entwickelte kognitive Systeme. Anwendungsgebiete sind die Sprach-, Text-, Gesichtserkennung uvm. So nehmen wir im folgenden Bild 3.2 aus der Anordnung eines Kreises, zweier Kreuze und eines Bogens ein Gesicht wahr.



Abbildung 3.2: Beispiel einer Erkennung eines Gesichts

Die Form des Bogens gibt uns zusätzlich die Information, dass das Gesicht Unzufriedenheit oder Trauer darstellt und dies relativ unabhängig vom Kulturkreis des Betrachters.

Seit den ersten Photographien im Jahr 1826 durch Joseph Nicéphore Niépce sind Bilder ein wesentlicher Bestandteil der Informationsweitergabe und -verarbeitung geworden. Mit der Einführung des Internets, welches 1969 aus einem Projekt der Advanced Research Project Agency des US-Verteidigungsministeriums entstand, haben die Möglichkeiten des Informationsaustauschs gerade von Bildern noch weiter zugenommen, genau wie der Wunsch, diese verarbeiten und auswerten zu können, sogar zu Überwachungszwecken [11]. Heutzutage leidet man nicht an einem Mangel an Informationen, sondern an einer schieren Informationsflut.

In der Medizin haben seit der Entdeckung der Röntgenstrahlen 1895 die bildgegebenen Verfahren wie Röntgenaufnahmen, Computertomographie, Szintigraphie, Sonographie, Magnetoresonanztomographie und diagnostische Thermographie immer mehr an Bedeutung in der Diagnostik gewonnen. Sie sollen dem Mediziner bereits ohne einen erheblichen Eingriff in den Körper ermöglichen, eine Einschätzung und die sich daraus ergebende Planung entsprechender Maßnahmen vorzunehmen. Ein Eingriff ohne eine entsprechende Einschätzung kann gerade im medizinischen Bereich unvorhersehbare Folgen haben und die Situation für den Patienten noch weiter verschlechtern [44].

Auch in anderen Bereichen versucht man die neuen Möglichkeiten der Bildverarbeitung zu nutzen, um weitere Einblicke in die entsprechenden Forschungsfelder zu gewinnen. In diesem Zusammenhang kann man Messdaten, wie die von elektronischen Nasen zur Bestimmung der Geruchsbelastung und Geruchssubstanz nennen [30]. Im Bereich der Kriminalistik und Kriminologie kann neben der Erkennung von Tätern [6] und Taten die Bildgebung das Vorgehen bei Straftaten [13] oder bei spezifischen Tätern veranschaulichen [82], wie beispielsweise das

---

bei sexuell motivierten Serienmördern [50].

Die Formanalyse setzt genau an dieser Informationsauswertung von Bildmaterial an. Ihr Ziel ist es, Objekte bzw. Objektgruppen trotz Transformationen (Rotation, Lage, Reflexion, etc.) wiederzuerkennen.

So ist in der folgenden Abbildung 3.3, die griechische Flagge trotz unterschiedlicher Lage (4,6), Größe (3), Drehung (1) und Spiegelung (3) als griechische Flagge zu sehen und wieder zu erkennen.

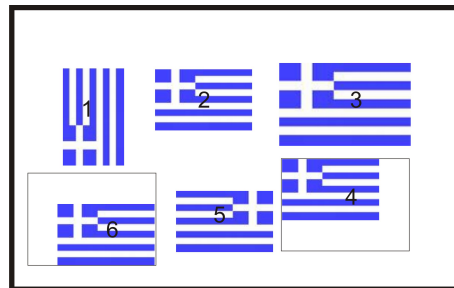


Abbildung 3.3: Beispiel für die unterschiedlichen Transformationen eines Objektes

Sobald man das Objekt im dreidimensionalen Raum betrachtet und dreidimensionale Transformationen berücksichtigen muss, sind aus den zweidimensionalen Bildern zunächst ein dreidimensionales Objekt zu bilden, um eine Vergleichbarkeit zu ermöglichen. Nur unter Beachtung der Bewegungen im dreidimensionalen Raum wird die Spinne aus der Ansicht von oben mit der Spinne auf Augenhöhe in Bild 3.4 vergleichbar.

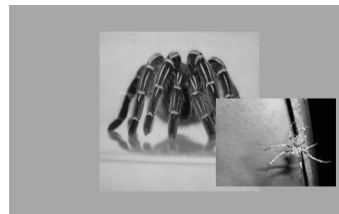


Abbildung 3.4: Beispiel für die notwendige Beachtung von Transformationen im dreidimensionalen Raum

Außerdem ist zu beachten, dass die Transformationen sich unterschiedlich auf die einzelnen Bereiche ein- und desselben Objektes auswirken können und damit die Wiedererkennung erschweren. So wird im Vergleich der beiden Spinnen im Bild 3.4 deutlich, dass die Proportion der Gliedmaße trotz derselben

---

Kategorie "Spinne" sehr unterschiedlich sein können.

Dies hängt davon ab, ob über die Formanalyse hinaus weitere Transformationen in der jeweiligen Anwendung für die Einordnung eines Objektes zu beachten sind. Sollte ein Parallelogramm in die gleiche Kategorie einzuordnen sein wie ein Rechteck, so ist die unterschiedliche Verschiebung der vier Punkte zu beachten. Eine solche zu vernachlässigende Transformation könnte sich aufgrund der Bewegungsabläufe des Objekts ergeben (Bewegung des Patienten während der Aufnahme, Atmung, Herzschlag, Mimik, Gestik etc.).

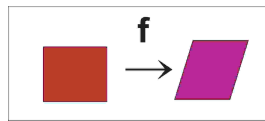


Abbildung 3.5: Transformation eines Rechtecks zu einem Parallelogramm

Die oberen zwei Punkte sind im Bild 3.5 weiter nach rechts verschoben worden als die unteren zwei Punkte.

Unter euklidischen Transformationen versteht man die Translation, die Spiegelung und die Drehung des gesamten Objektes. Im obigen Fall handelt es sich also um eine nicht-euklidische Transformation.

Wenn man Transformationen eliminiert, geht man davon aus, dass die Informationen in den Transformationen unbeachtet bleiben können. Dies ist nicht immer der Fall. Die beiden Verkehrsschilder in Bild 3.6 haben einen unterschiedlichen Sinn und dürfen also nicht mit Hilfe einer geometrischen Transformation als gleich angesehen werden.

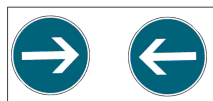


Abbildung 3.6: Beispiel: Relevante Informationen in der Reflektion

## 3.2 Eine kurze Historie der Formanalyse

Interesse an der Formanalyse begann 1977. D.G. Kendall [71] stellte einerseits in einem Artikel eine Möglichkeit vor, Formen als Elemente des Komplexen darzustellen. Er benutzte dabei die Äquivalenz  $\mathbb{R}^2 \equiv \mathbb{C}$ . Zweidimensionale Punkte  $(x, y)$  werden im Komplexen dann als  $z = x + iy$  abgebildet.

K.V. Mardia [76] erforschte andererseits die Verteilung der Formen von Dreiecken. Wenn alle Landmarken unabhängig und identisch nach einer rotationsymmetrischen Verteilung verteilt sind, dann ergibt sich eine Gleichverteilung im betrachteten Raum der Formen [69]. In [69] sind die unterschiedlichen Räume für die zu betrachtenden Formen definiert<sup>1</sup>.

Die theoretischen Ansätze, welche eine Menge an interessanten Forschungsfelder (Dinosaurierspuren, Siedlungen in der späten Bronzezeit in England etc.) umfasst, sind von D. Kendall [72] und F. Bookstein [7] publiziert worden. Eine detaillierte Darstellung, sowie der weiteren Entwicklung kann man in den Büchern von C.G. Small [83] und I.L. Dryden & K.V. Mardia [60] finden<sup>2</sup>.

In der Pilotstudie von Giebel [32] wird erstmalig bei den Nierentumoren hinsichtlich der auf den dreidimensionalen Schwerpunkt zentrierten zweidimensionalen Objekte von der Gestalt gesprochen. Werden diese zunächst normiert, wird von Form I gesprochen, bei einer darauf folgenden Zentrierung auf den zweidimensionalen Schwerpunkt von Form II. Unter Originaldaten werden die bereits unter Beachtung des dreidimensionalen Schwerpunktes erfassten Daten verstanden<sup>3</sup>.

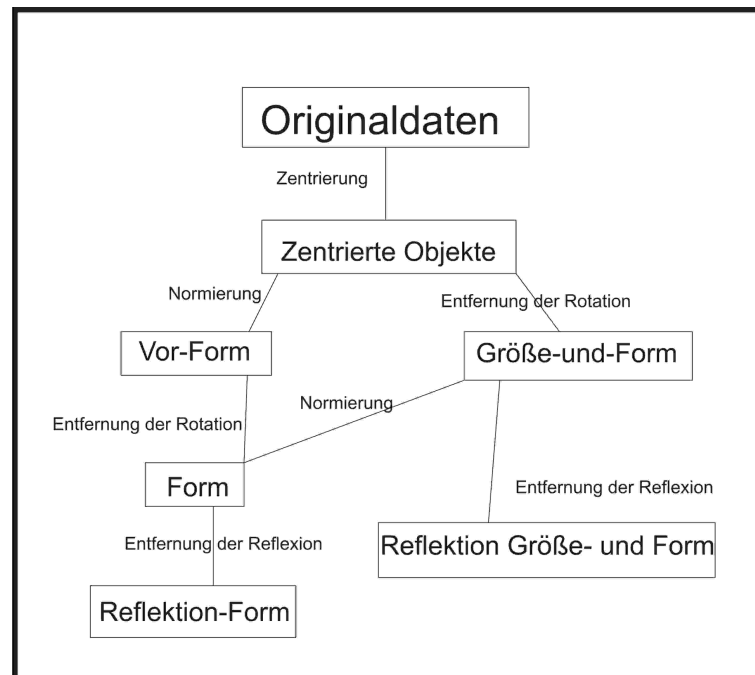


Abbildung 3.7: Definition der Räume nach Mardia & Dryden

<sup>1</sup> siehe S. 58 [69]

<sup>2</sup> In [83] ist der Überblick auf S.199/200 zu finden und bei [60] auf den S.9-21

<sup>3</sup> Eine tabellarische Übersicht findet sich in [38] auf S.47

---

Die Form einer Menge unabhängiger Teilchen diffundierend gemäß der Brownschen Bewegung kann nach [71] wiederum als eine Brownsche Bewegung im Komplexen verstanden werden.

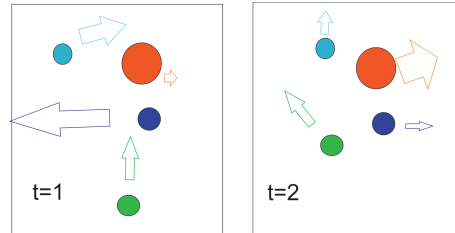


Abbildung 3.8: Beispiel für die Brownsche Bewegung zum Zeitpunkt  $t = 1$  und  $t = 2$

Aufbauend auf den Arbeiten von Kendall, nutzt H. Ziezold [90] das Umschreiben der zweidimensionalen Punkte zu Vereinfachungen bei der Beweisführung und Berechnung [90].

Um die Brownsche Bewegung mathematisch zu verstehen, braucht man den Begriff der Filtrierung.

**Definition 3.2.1.** Sei  $T$  eine beliebig geordnete Indexmenge und  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Eine Filtrierung ist eine Familie  $(\mathcal{F}_t), t \in T$ , von  $\sigma$ -Algebren, bei der  $(\mathcal{F}_t), t \in T$ , aufsteigend geordnet ist, das heißt  $\forall s, t \in T, s < t : \mathcal{F}_s \subseteq \mathcal{F}_t$ . Im speziellen wird dadurch die zu dem Zeitpunkt  $t$  vorhandene Information über den Verlauf des zufälligen Prozesses modelliert.

**Beispiel 1.** Bei der Betrachtung der Aktienkurse in [36] wird das mit dem stochastischen Prozess kombinierte Neuronale Netz zur Vorhersage von Aktienkursen stets unter der Kenntnis eines bestimmten bereits vergangenen Zeitabschnitts kallibriert. Die Ordnung der vorhandenen Information ergibt sich durch die Zeit. Die zukünftigen Werte werden mit Hilfe der Information aus den vergangenen Werten vorhergesagt.

**Definition 3.2.2.** Die Brownsche Bewegung oder Wiener-Prozess ist ein stochastischer Prozess  $(W_t)_{t \in \mathbb{R}_+}$  auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ , der an die Filtrierung  $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$  angepasst ist und folgende Bedingungen erfüllt:

1.  $W_0 = 0$  ( $P$ -fast sicher). Das bedeutet, dass es nahezu sicher ist, dass der Startwert des Wiener-Prozesses 0 ist.
2. Für ein gegebenes  $s$  mit  $t > s \geq 0$  sind alle Inkremente  $W_t - W_s$  stochastisch unabhängig von  $\mathcal{F}_s$ . Der Wiener-Prozess hat also unabhängige Zuwächse.

- 
3.  $W_t - W_s \sim \mathcal{N}(0, t - s) \quad \forall 0 \leq s < t$ . Die Zuwächse sind also stationär und normalverteilt mit Mittelwert 0 und Varianz  $t - s$ .
  4. Die einzelnen Pfade sind (P-)fast sicher stetig.

**Beispiel 2.** In [36] wird bei der stochastischen Modellierung angenommen, dass die Veränderung des Kurses zu jedem Zeitpunkt von den in der Vergangenheit vorhandenen Veränderungen unabhängig ist. Das Neuronale Netz setzt keine Unabhängigkeit voraus und bestimmt den weiteren Kursverlauf durch eine entsprechend unterschiedliche Gewichtung der Zeitpunkte bei der Berechnung des Mittelwertes und der Varianz.

**Definition 3.2.3.** Ein Martingal ist ein stochastischer Prozess, in dem der Erwartungswert einer Beobachtung dem Wert der vorherigen Beobachtung entspricht. Hinsichtlich der Form erwarten wir zukünftig die bereits beobachtete Form. Nach Lévy kann ein Wiener-Prozess  $(W_t)_{t \in \mathbb{R}_+}$  durch folgende zwei Eigenschaften charakterisiert werden:

- $W_t$  ist ein stetiges lokales Martingal mit  $W_0 = 0$
- $W_t^2 - t$  ist ein Martingal

**Beispiel 3.** Bezogen auf die Aktienkurse geht man bei einem Martingal davon aus, dass der zukünftige zu erwartende Wert der Aktie gleich dem vorherigen beobachteten Wert entspricht.

Übertragen auf die Formen ist die letzte beobachtete Form zu erwarten. Trotz der Vielzahl an Anwendungsmöglichkeiten der statistischen Formanalyse fristet sie außerhalb der Mathematik noch ein Nischendasein und selbst in der Mathematik gibt es nur sehr wenige, die mit diesem Forschungsgebiet vertraut sind.

### 3.3 Die Grundbegriffe der statistischen Formanalyse

Die statistische Formanalyse ist ein Verfahren, welches erlaubt, aus dem Umriss von Objekten eine mittlere zu erwartende Form oder Gestalt einer Menge von Objekten zu errechnen.

Um die Form eines Objektes zu ermitteln, benötigt man klar definierte Punkte zur Erfassung der geometrischen Information.

**Definition 3.3.1.** Eine **Landmarke** ist ein Punkt, der bei jedem Objekt in der zu untersuchenden Objektmenge vergleichbar erhoben wird. Es gibt durchnummerierte und nicht durchnummerierte Landmarken. Bei durchnummerierten Landmarken kann die Relevanz bestimmter Bereiche des Objektes näher untersucht werden.

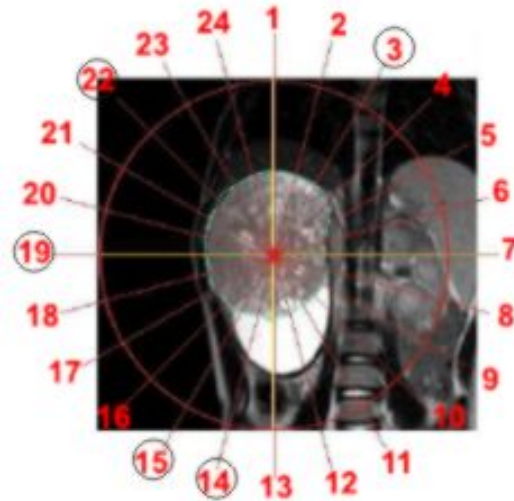


Abbildung 3.9: Erhebung von 24 Landmarken um den Tumor und Hervorhebung der in der Pilotstudie fünf relevanten Landmarken [32]

Die Erhebung der Landmarken und damit die Gewichtung bestimmter Bereiche des Objektes liegt gänzlich in der Hand des Anwenders. Die Wahl der Landmarken kann rein explorativ erfolgen oder sich aus theoretischen Annahmen und Vorüberlegung des jeweiligen Fachgebietes ergeben, wie im Falle des Gehirns (Bild 3.10).

Zusätzlich zu den Landmarken kann als weitere geometrische Information die Verbindungen zwischen den Landmarken erhoben werden. Für  $k$  Landmarken kann man bis auf maximal  $\binom{k}{2}$  Verbindungen kommen, wenn jede Landmarke mit allen anderen Landmarken verbunden wird.

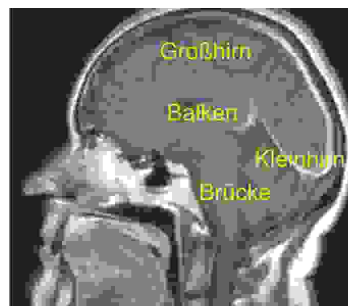


Abbildung 3.10: Aufnahme des Gehirns und die medizinische Bedeutung einiger Areale

---

**Bemerkung 1.** Die Relevanz bestimmter Bereiche des Objektes, sowie die Vergleichbarkeit der Objekte untereinander hängen allein von der Erhebung (Landmarken) ab.

Aus den erhobenen geometrischen Informationen muss nun eine charakteristische Form der Objekte gewonnen werden. Die zur Differenzierung wichtigen Eigenschaften und damit die eventuelle Notwendigkeit der Eliminierung der Lage, Größe und der Rotation hängt von der Anwendung ab. Es stellt sich für den Anwender die Frage, ob die Eliminierung der jeweiligen Information eine Verbesserung der Differenzierung der Objektgruppen zur Folge hat.

**Definition 3.3.2.** Die **Gestalt** umfasst sämtliche geometrische Informationen nach dem Entfernen der Unterschiede hinsichtlich der Lage des Objektes

**Definition 3.3.3.** Die **Vorform** umfasst sämtliche geometrische Informationen nach dem Entfernen der Lage und Größe des Objektes

**Definition 3.3.4.** Die **Form** umfasst sämtliche geometrische Informationen nach dem Entfernen der Lage, Größe und Rotationseffekte des Objektes

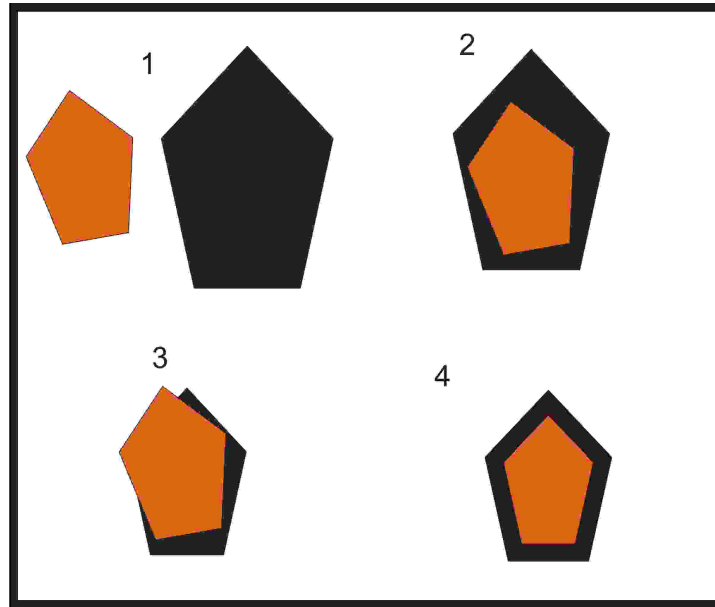


Abbildung 3.11: Vorgehen der Formanalyse: 1 Erhebung, 2 Gestalt, 3 Vorform und 4 Form

Nachdem die Daten erhoben sind, werden  $k$  *Landmarken* vom Rand des Objektes  $o$  gewählt:  $l_j \in \partial o$  für  $j = 1, \dots, k$ . Die Landmarke hat die Dimension  $m$  des



---

Raumes  $V$ , in welchem die Umrandung des Objektes liegt:  $\partial o \subset V$ .

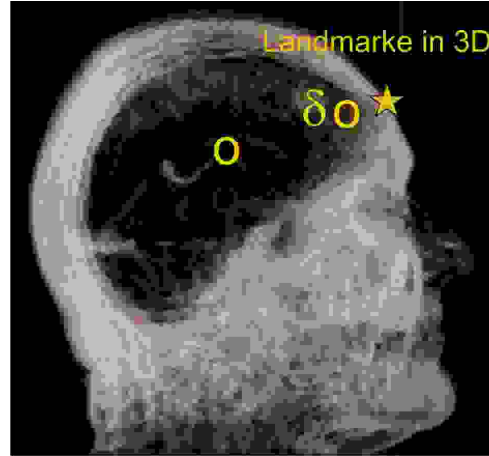


Abbildung 3.12: Landmarke im dreidimensionalen Raum am Rand des Schädels eines Patienten

Im Fall von medizinischen Bildern liegt die Umrandung  $\partial o$  im zwei- oder dreidimensionalen Raum, demnach  $m = 2$  oder  $m = 3$  bzw.  $\partial o \subset \mathbb{R}^2$  oder  $\partial o \subset \mathbb{R}^3$ . Anwendungen in höherdimensionalen Räumen wäre bei Beachtung weiterer Kenngrößen der Landmarken wie bspw. des Zeitpunkts der Aufnahme, der gemessenen Gewebedichte oder Durchblutung denkbar. Durch das Hinzunehmen der Zeit bekommen wir zum Beispiel vierdimensionale Landmarken in den Anwendungen der Herzfrequenz und dem Vorgehen von Serienmördern.

Da der Zeitachse meist in der Anwendung eine besondere Bedeutung zukommt (Diastole und Systole, Mord Nr. 1... Mord Nr. k) und teilweise nicht kontinuierlich gemessen wird bzw. gemessen werden kann, führt man die Formanalyse dabei zunächst im  $\mathbb{R}^3$  durch. Die Vergleichbarkeit der jeweiligen Momentaufnahme ist nur unter Beachtung der jeweiligen Anwendung und der Deutung des Zeitpunkts möglich.

Jedes der  $n$  analysierten Objekte wird also durch  $k$  Landmarken im  $m$ -dimensionalen Raum beschrieben, das heißt, man braucht  $k \cdot m$  Koordinaten für jedes Objekt.

$$\forall i = 1 \dots n, \forall j = 1 \dots k, o_i = \{l_1 \dots l_k\}, l_j \in \mathbb{R}^m, o_i \in \mathbb{R}^{mk} \quad (3.1)$$

Die bereits definierte *Gestalt* des Objektes erhält man dann durch eine vergleichbarer Lage der Objekte und damit der Landmarken entweder durch *Zentrierung* oder durch eine entsprechende Art der Erhebung der Daten (Erfassung

---

der Daten ausgehend vom dreidimensionalen Schwerpunkt etc.).

Die *Zentrierung der Landmarken* wird folgendermaßen erreicht:

1. Man bestimmt das Zentrum  $z^i$  der  $k$  Landmarken des  $i$ -ten Objektes  $o_i$  durch den Mittelwert

$$z^i = \frac{1}{k} \sum_{j=1}^k l_j^i \quad (3.2)$$

2. Man zentriert das Objekt  $o_i$  auf dieses Zentrum, indem jede Landmarke um das bereits berechnete Zentrum  $z^i$  verschoben wird

$$l_i^* = l_j^i - z^i \quad (3.3)$$

**Beispiel 4. Zentrierung im zweidimensionalen Raum** Sei das Dreieck mit den Eckpunkten  $A(5|5), B(5|1), C(8|3)$

Da  $\mathbb{R}^2 \equiv \mathbb{C}$  kann man die Berechnungen im komplexen Raum durchführen:  $z = \frac{1}{3}(5 + 5i + 5 + 1i + 8 + 3i) = \frac{1}{3}(18 + 9i) = 6 + 3i$ . Das Zentrum des Dreiecks liegt somit im Punkt  $Z(6|3)$ . Das Zentrum wird nach der Verschiebung in den Ursprung  $Z'(0|0)$  verlegt, indem die erste Koordinate jeder Landmarke um 6 Einheiten und die der zweiten Koordinate um 3 Einheiten verschoben wird  $A^z(-1|2), B^z(-1|-2), C^z(2|0)$ .

**Beispiel 5. Zentrierung im dreidimensionalen Raum** Sei das Dreieck mit den Eckpunkten  $A(5|5|1), B(5|1|0), C(8|3|2)$ .

Die Verschiebung für die x-Achse ergibt sich aus  $\frac{1}{3}(5+5+8) = 6$ , für die y-Achse aus  $\frac{1}{3}(5+1+3) = 3$  und für die z-Achse aus  $\frac{1}{3}(1+0+2) = 1$ . Das zentrierte Dreieck hat also die Koordinate  $A^z(-1|2|0), B^z(-1|-2|-1), C^z(2|0|1)$ .

Eine Alternative für die so vorgenommene Zentrierung ist die Verwendung von *Bookstein-Koordinaten* [7] oder durch die Multiplikation mit der sogenannten *Helmert-Matrix*. Problem bei der Helmert-Matrix ist, dass bei der Zentrierung eine Landmarke verloren geht<sup>4</sup> und das so erhaltene Objekt nicht mehr mit dem ursprünglichen Objekt vergleichbar ist.

Bei Bookstein-Koordinaten wird eine "Basis" (Linie zwischen zwei Punkten) benötigt. Die Projektion dieser Basis auf vorbestimmte Punkte wird für alle weiteren Punkte durchgeführt. Der Nachteil der Bookstein-Koordinaten ist, dass die Varianz der Punkte, die die Basis bilden, eliminiert wird und damit ist die Varianz dieser Punkte nicht mehr vergleichbar mit der der anderen Punkte.

Die bereits definierte *Vorform* des Objektes wird dann durch die *Normierung* erhalten. Durch die Normierung, wird die Größe der Objekte neutralisiert. In

---

<sup>4</sup> siehe [60], S.60

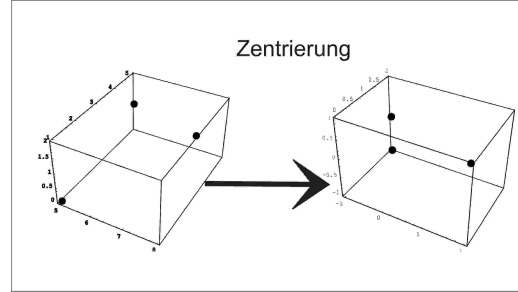


Abbildung 3.13: Zentrierung im dreidimensionalen Raum gemäß dem obigen Beispiel 5

Anwendungen, in denen keine Rotation der Objekte möglich ist, erhält man durch Zentrierung und Normierung bereits die *Form* des Objektes.

In der vorliegenden Untersuchung wird für die *Normierung der Landmarken* die euklidische Norm verwendet:

1. Zunächst wird die euklidische Norm für das  $i$ -te Objekt berechnet :

$$\|o_i\| = \sqrt{\sum_{j=1}^m |l_j^i|^2} \quad (3.4)$$

2. Jede der Landmarken wird dann durch den Wert der Normierung geteilt:

$$l_i^0 = \frac{l_i^*}{\|o_i\|} \quad (3.5)$$

**Beispiel 6. Normierung im zweidimensionalen Raum** Sei das in Beispiel 4 verwendete Dreieck. Es ergibt sich:

$$\|o_i\| = \sqrt{|(5 + 5i)|^2 + |(5 + 1i)|^2 + |(8 + 3i)|^2} = \sqrt{50 + 26 + 73}.$$

Das zentrierte und normierte Dreieck hat also folgende Koordinaten:

$$A^n(\frac{5}{\sqrt{149}} | \frac{5}{\sqrt{149}}), B^n(\frac{5}{\sqrt{149}} | \frac{1}{\sqrt{149}}), C^n(\frac{8}{\sqrt{149}} | \frac{3}{\sqrt{149}}).$$

**Beispiel 7. Normierung im dreidimensionalen Raum** Sei eine Kugel, beschrieben durch ihren Mittelpunkt  $A(0|0|0)$  und einem Punkt an der Kugeloberfläche  $B(2|2|2)$ . Die Norm der Kugel ergibt dann:

---


$$\|o_i\| = \sqrt{(0^2 + 0^2 + 0^2) + (2^2 + 2^2 + 2^2)} = \sqrt{12}.$$

Die normierte Kugel ist also durch den Mittelpunkt  $A^n(0|0|0)$  und den Punkt  $B^n(\frac{2}{\sqrt{12}}|\frac{2}{\sqrt{12}}|\frac{2}{\sqrt{12}})$  definiert.

Normierung

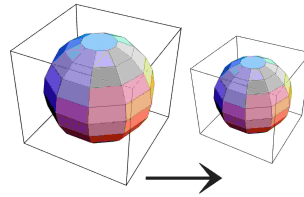


Abbildung 3.14: Normierung im dreidimensionalen Raum gemäß Beispiel 7

Durch die Normierung wird sicher gestellt, dass alle Objekte auf “gleiche Größe” gebracht werden und die geometrische Information der Objekte ohne Beachtung ihrer Lage und Größe nach diesen beiden Schritten miteinander verglichen werden können.

**Bemerkung 2.** Will man Objekte mit unterschiedlicher Anzahl an Landmarken miteinander vergleichen, ist eine Multiplikation bei jeweils  $k$  Landmarken des Objektes mit  $\frac{1}{k}$  bei der Berechnung der Distanz notwendig.

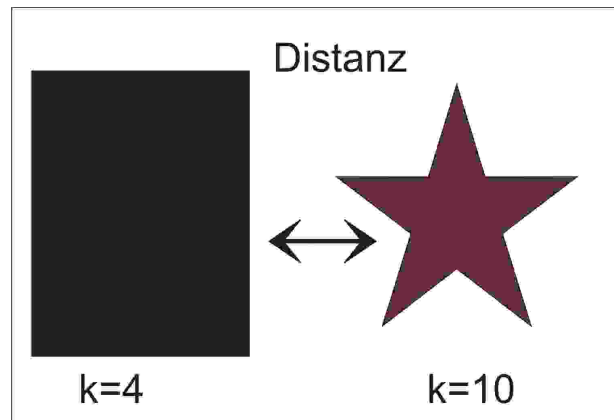


Abbildung 3.15: Berechnung der Distanz bei unterschiedlicher Anzahl von Landmarken

**Beispiel 8.** Dieses Problem ergibt sich zum Beispiel beim Vergleich von Schriftzeichen. Im folgenden werden die zwei chinesischen Schriftzeichen shí (zehn) und

xià (nächste) verglichen. Als Landmarken werden die Schnittpunkte und Endpunkte des Buchstabens im Fließtext gewählt:

$\mathbf{p} : A(27, 926|221, 970), B(0, 589|197, 220), C(57, 188|197, 220), D(27, 355|161, 782), E(28, 716|197, 220)$  und

$\mathbf{q} : A'(66, 739|212, 407), B'(126, 230|212, 918), C'(121, 235, 184, 845), D'(94, 293|160, 095), E'(94, 293|212, 407), F'(93, 750|219, 375)$ .

Für die Normierung ergibt sich ohne die Beachtung von  $k$  für das erste eine Norm von 500.133 und für das zweite eine von 495.736. Unter Beachtung von  $k$  eine Norm von 83.3554 und 99.1472. Um einen Abstand der dann zu zentrierenden Objekte bei unterschiedlicher Anzahl an Landmarken  $k$  zu erhalten wird die Differenz der Norm als Abstand:  $d = \|\frac{1}{5}\mathbf{p}\| - \|\frac{1}{6}\mathbf{q}\| = 0.0500973$  verwendet. Da bei Schriftzeichen keine Rotation vorliegt, kann das Alphabet nun in eine Rangfolge überführt werden und die Erkennung eines Buchstabens über den Abstand zu anderen Buchstaben erfolgen.

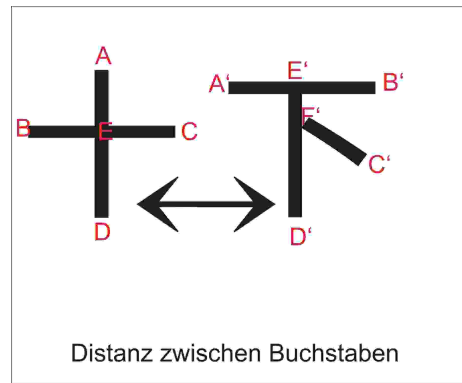


Abbildung 3.16: Berechnung der Distanz bei unterschiedlicher Anzahl von Landmarken

**Beispiel 9.** Bei den 2010 untersuchten  $n = 16$  Raubmördern [51] ergibt sich die unterschiedliche Anzahl an Landmarken durch die Anzahl an begangenen Taten. Für den Kriminologen stellt sich die Frage, ob trotz der unterschiedlichen Anzahl, eine Vergleichbarkeit der Täter hinsichtlich ihres räumlichen Profils möglich ist. Die 35jährige Irmgard Swinka begang 14 Raubmorde in den Nachkriegsjahren und wurde 1949 verurteilt. Eckhard Krüger ist ein ebenfalls bundesweit agierender Raubmörder. Ihm konnten vier Raubmorde nachgewiesen werden.

Die Norm für die Täterin Irmgard Swinka beträgt 14.072 und die Norm für den Täter Eckhard Krüger 26.5607. Der Abstand der dann zentrierten Profile beträgt  $-0.515032$ . Trotz der unterschiedlichen Anzahl an Taten können die beiden Raubmörder miteinander verglichen werden. Ein entsprechend ähnliches

räumliches Vorgehen kann Rückschlüsse zur Mobilität und zu Charakteristika des Täters erlauben.

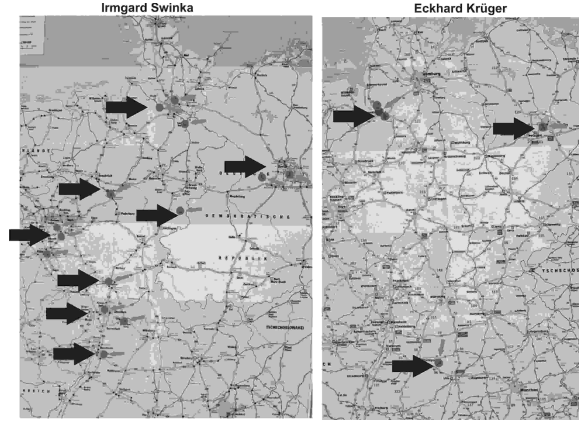


Abbildung 3.17: Zwei Raubmörder im Vergleich, Legende: Pfeile kennzeichnen die Orte, in denen die Taten verübt worden, Kreise den genauen Tatort

Nach dem Zentrieren und Normieren der Objekte stellt sich die Frage, ob die Objekte nun in der optimalen Lage zueinander liegen, oder ob die Beachtung der Rotation möglicherweise die Distanz zwischen ihnen noch verringern kann.

**Definition 3.3.5.** Sei  $u \in SO(m)$  eine Drehung und  $a \in \mathbb{R}^m$  die Verschiebung des Objektes, so wird in  $\mathbb{R}^{mk}$  eine Distanz  $\delta$  definiert durch:

$$\delta(x, y) = \inf_{u \in SO(m), a \in \mathbb{R}^m} \|x - uy - a\|, \forall u, v \in SO(m) \times \dots SO(m), \forall a, b \in \mathbb{R}^{mk} \text{ für } x, y \in \mathbb{R}^{mk}.$$

Die beiden Objekte  $x$  und  $y$  werden dabei derart in ihrer Lage verschoben und gedreht, dass sie den geringsten Abstand zueinander aufweisen. Es gilt:

$$\delta(ux + a1, vy + b1) = \delta(x, y) \text{ für } x \in \tilde{x}, y \in \tilde{y}.$$

$\tilde{x}$  und  $\tilde{y}$  sind dabei die Äquivalenzklassen von  $x$  und  $y$ .  $x$  und  $y \in \mathbb{R}^{mk}$  gehören zu der  $k$ -dimensionalen Äquivalenzklasse in  $\mathbb{R}^m$  modulo der euklidischen Bewegungen. Der nun betrachtete metrische Raum aller Objekte ist der der Quotientenmetrik. Im Weiteren wird dieser mit  $(F^{mk}, d)$  bezeichnet. Für  $x, y \in \mathbb{R}^{mk}$  hat man  $\tilde{x}, \tilde{y} \in F^{mk}$  und

$$d(\tilde{x}, \tilde{y}) = \delta(x, y). \quad (3.6)$$

Von daher kennzeichnet  $\sim$  im weiteren die entsprechende Äquivalenzklasse in dem betrachteten Raum. Egal in welcher Art die beiden Objekte nun verschoben und gedreht werden, der so bestimmte Abstand der beiden Objekte bleibt davon unberührt.

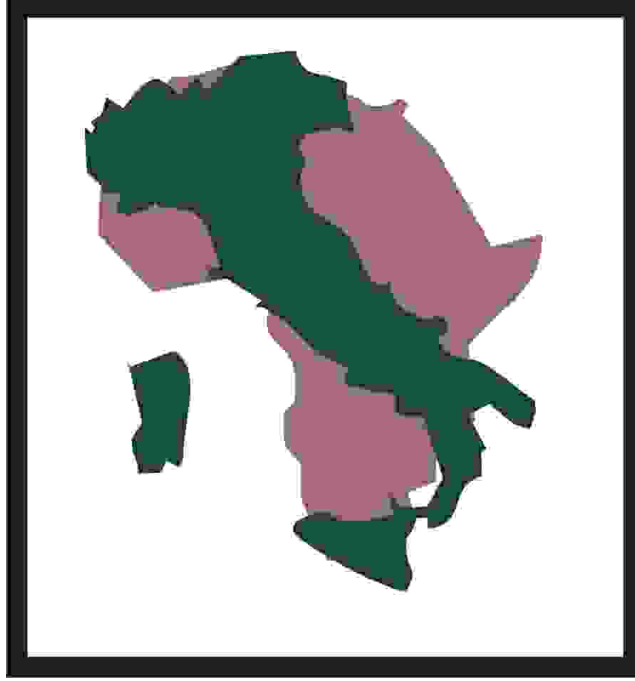


Abbildung 3.18: Afrika und Italien in optimaler Lage zueinander

**Beispiel 10.** Das Dreieck  $\mathfrak{x}$   $A(0|1), B(0|0), C(3|0)$  wird um eine Einheit auf der ersten Achse nach rechts verschoben und um eine Einheit nach unten auf der zweiten Achse. Es ergibt sich das Dreieck  $\mathfrak{y}$   $A(1|0), B(1|-1), C(4|-1)$ . Beide Dreiecke sind nach obiger Definition als äquivalent anzusehen. Die Distanz beträgt  $\delta(\mathfrak{x}, \mathfrak{y}) = 0$ .

**Bemerkung 3.** Nur die euklidischen Bewegungen finden dabei Beachtung. Um nicht-euklidische Bewegungen zu beachten, ist der Ansatz entsprechend zu erweitern:  $\delta'(x, y) = \inf_{u \in SO(m), a \in \mathbb{R}^m} \|Ax - uBy - a\|$  mit einer  $m \times k$  Matrix  $A$  und  $B$ .

**Definition 3.3.6.**  $x_i$  und  $x_j \in \mathbb{R}^{mk}$  befinden sich in **optimaler Lage** zueinander, wenn der Abstand der beiden Objekte nach Drehung und Verschiebung minimal ist. Die Drehungen  $u_i \in \mathbb{R}^{m \times m} \times \dots \mathbb{R}^{m \times m} \subset \mathbb{R}^{mk \times mk}$  sind in einer Untergruppe von  $SO(m)$  der umkehrbaren homogenen linearen Abbildungen  $GL(m)$  im  $\mathbb{R}^m$ , sie sind orientierungstreu und längenerhaltend, und für alle  $a_1 \dots a_n \in \mathbb{R}^m$  gilt

$$\sum_{1 \leq i < j \leq n} |x_i - x_j|^2 = \inf \sum_{1 \leq i < j \leq n} |u_i x_i + a_i 1 - u_j x_j - a_j 1|^2 \quad (3.7)$$

Die  $k$ -eckigen Objekte  $x_1$  und  $x_2$  liegen bei Erfüllung obiger Gleichung in optimaler Lage zueinander.

---

Nachdem die Objekte zentriert und normiert sind und wir davon ausgehen können, dass sie in optimaler Lage zueinander stehen, stellt sich die Frage, ob wir anhand der nun noch vorliegenden geometrischen Informationen Gruppen von Objekten bilden können. Da die Gruppe in den untersuchten Anwendungsfällen meist vorgegeben sind, bietet es sich an, die Verteilung der bereits definierten Gruppen mittels eines Erwartungswerts oder der "mittleren Form", sowie die Abweichung von der mittleren Form mittels der Varianz zu beschreiben und diese Kennwerte für eine Differenzierung der Gruppen zu verwenden.

### 3.4 Die mittlere Form

Die mittlere Form oder Gestalt einer Gruppe von Objekten soll diese Objekte optimal charakterisieren, unabhängig von ihrer ursprünglichen Lage und Größe. Die mittlere Form ist der Erwartungswert nach Fréchet der Menge der zufälligen Objekte.

**Definition 3.4.1. Erwartungswert nach Fréchet [62]:** Sei  $X$  eine Zufallsvariable von einem Wahrscheinlichkeitsraum  $(\Omega, \mathbf{F}, \mathbf{P})$  in einen metrischen Raum  $(\Xi, d)$ . Ein Element  $\mu \in \Xi$  ist der Erwartungswert von  $X$  bei  $k$  Objekten, wenn

$$\sum_{j=1}^k d(x_j, \mu)^2 = \inf_{\alpha \in \Xi} \sum_{j=1}^k d(x_j, \alpha)^2, \forall \alpha \in \Xi.$$

Der Erwartungswert  $\mu$  weist also von allen Elementen im metrischen Raum den niedrigsten Abstand zu der Zufallsvariablen  $X$  auf.

Auf unsere Notationen übertragen, ergibt sich also die folgende Definition der mittleren Form.

**Definition 3.4.2.** Die **mittlere Form** einer Menge von Objekten  $o_i, i = 1, \dots, n$  ist das Objekt  $m$  definiert durch:

$$\sum_{i=1}^n d(o_i, \mathbf{m})^2 = \inf_{\alpha \in \mathbb{R}^{mk}} \sum_{i=1}^n d(o_i, \alpha)^2 \quad (3.8)$$

Vorausgesetzt ist die Endlichkeit des obigen Ausdrucks. Die mittlere Form  $\mathbf{m}$  entspricht dem Erwartungswert nach Fréchet bei zufälligen Objekten. Es stellt sich nun die Frage, ob das Infimum immer von mindestens einem der Objekte erreicht wird:

**Bemerkung 4.** Stetigkeit einer Funktion führt dazu, dass keine Sprünge auftreten. Der Zwischenwertsatz sagt aus, dass alle von der Funktion abgebildeten



---

Werte des Intervalls  $[a, b]$ , nämlich  $[f(a), f(b)]$  mindestens einmal erreicht wird:  
Da für die Metrik die Dreiecksungleichung gilt, erhält man

$$\begin{aligned} d(x, y) &\leq d(x, x_0) + d(x_0, y) \\ -d(x, x_0) &\leq d(x_0, y) - d(x, y) \\ d(x_0, y) - d(x, x_0) &\leq -d(x, y) \\ |d(x_0, y) - d(x, x_0)| &\leq d(x, y) \\ \delta = |d(x_0, y) - d(x, x_0)| &\leq d(x, y) \leq \epsilon. \end{aligned}$$

**Bemerkung 5.** Es gibt immer eine untere Schranke. Diese wird erreicht und für eine Menge von Formen ergibt sich also mindestens eine mittlere Form

Je nach Metrik kann es mehrere unterschiedliche mittleren Formen geben.

**Beispiel 11.** Sei die diskrete Metrik

$$d(x, y) = 1, \text{ wenn } x \neq y \text{ und } d(x, y) = 0, \text{ wenn } x = y$$

Da bei dieser Metrik die Distanz zwischen verschiedenen Objekten immer den gleichen Wert ergibt, kann jedes Element einer Menge unterschiedlicher Formen auch deren mittlere Form darstellen.

**Beispiel 12.** Bei ehemals inhaftierten Straftätern wird die erneute Straffälligkeit, der Erhalt der Disziplinarmaßnahme und die Partnerschaft mit Ausprägung 1/"ja" und 0/"nein" erfasst. Es gibt also acht verschiedene Profile: "111" ... "000". Die mittlere Form ist das Profil, welches am häufigsten vorkommt und entspricht damit dem Modus der erhobenen Daten. Es stellt das Infimum dar.

**Beispiel 13.** Die Dreiecke  $\mathfrak{g} : A(1|0|0), B(1|1|0), C(0|1|4)$  und  $\mathfrak{h} : A(1|0|0), B(1|2|0), C(0|1|5)$  haben in der diskreten Metrik den Abstand  $d(\mathfrak{g}, \mathfrak{h}) = 2$ .

Nach der Zentrierung  $\mathfrak{g}^z : A(\frac{1}{3}|\frac{2}{3}|\frac{4}{3}), B(\frac{1}{3}|\frac{1}{3}|\frac{4}{3}), C(-\frac{2}{3}|\frac{1}{3}|\frac{8}{3})$  und  $\mathfrak{h}^z : A(\frac{1}{3}|-1|\frac{5}{3}), B(\frac{1}{3}|1|\frac{5}{3}), C(-\frac{2}{3}|0|\frac{10}{3})$  haben sie den Abstand  $d(\mathfrak{g}^z, \mathfrak{h}^z) = 3$ .

## Zentrierung

Ziezold definierte die Zentrierung und die mittlere Form für die Euklidische Metrik  $||\cdot||$  [90]. Er blieb in seinen Anwendungen im zweidimensionalen Raum. Erst in dieser Arbeit werden seine Ansätze und Vorüberlegungen auf den dreidimensionalen Raum übertragen.

Für die von der p-Norm  $|\cdot|_p$  induzierten Metrik  $d_p(x, y) := |x - y|_p$  wird angenommen, dass für die Varianz gilt:

$$s^2 = \frac{1}{n^2} \sum_{1 \leq i < j \leq n} |x^i - x^j|^2 = \frac{1}{n^2} \sum_{1 \leq i < j \leq n} |u_i \cdot x_i + a_i 1 - u_j \cdot x_j + a_j 1|^2. \quad (3.9)$$

---

Erinnern wir daran, dass  $m$  die Dimension des Raumes ist, in dem wir die Objekte beobachten,  $k$  die Anzahl an Landmarken pro Objekt und  $n$  die Anzahl an Objekten. Sei die Abbildung  $g : \mathbb{R}^{mk} \rightarrow \mathbb{R}_0^+$

$$g(a) : |x-a|_p = \sqrt[p]{\sum_i^k \sum_j^m (x_{ij} - a_{ij})^p} = \sqrt[p]{\sum_i^k \sum_j^m (x_{ij} - a_j)^p} = \sqrt[p]{\sum_i^k |x_i - a|_p^p}.$$

Eine Differenzierung von  $g$  nach  $a_l$ , für  $l = 1, \dots, m \leq k$  ergibt :

$$\frac{\partial g}{\partial a_l} = - \sum_{i=1}^k (x_{il} - a_l)^{p-1} \sqrt[p]{\sum_{i=1}^k |x_i - a|_p^p} \quad (3.10)$$

Das zentrierte Objekt ergibt sich dann als Extremalstelle von  $g$ . Die Zentrierung erfolgt je nach Objekt unterschiedlich. Erst nach der Zentrierung sind die Objekte alle in der gleichen Lage. Für  $p = 2$  ergibt sich nach Differentiation von  $g(a) = \|x - a\|$  :

$$a_l = \frac{1}{k} \sum_{j=1}^k x_j \in \mathbb{R}^m. \quad (3.11)$$

### Mittlere Form

Sei die Abbildung  $f : \mathbb{R}^{mk} \rightarrow \mathbb{R}_0^+$ . Die mittlere Form  $m$  für die  $p$ -Norm ergibt sich aus:

$$f(m) := \sum_r^n \|x_i - m\|_p^2 = \sum_{r=1}^n \left( \sum_{ij}^{km} |x_{r,i,j} - m_{i,j}|_p^p \right)^{\frac{2}{p}} \quad (3.12)$$

Differentiation nach  $m$ :

$$\frac{\partial f}{\partial m_{pq}} = -2 \sum_{r=1}^n (|m - x_r|_p^{\frac{2}{p}-1} (x_{r,pq} - m_{pq})^{p-1}) \quad (3.13)$$

Für den Spezialfall  $p = 2$  nehmen wir  $|x_i| = \sqrt{(x_i)^2}$

$$m = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.14)$$

Für  $p \neq 2$  sind die Gleichungen  $df/dm = 0$  und die Lösungen für  $m$  jeweils neu zu bestimmen. Um zu zeigen, dass es ein Minimum ist, ist das Ergebnis in  $f''(m)$  einzusetzen. Da  $f''(m) > 0$ , handelt es sich um ein Minimum.

---

## Rotationen

Im Weiteren sind die Rotationen ein- und derselben Form wie im unteren Bild 3.19 zu beachten:

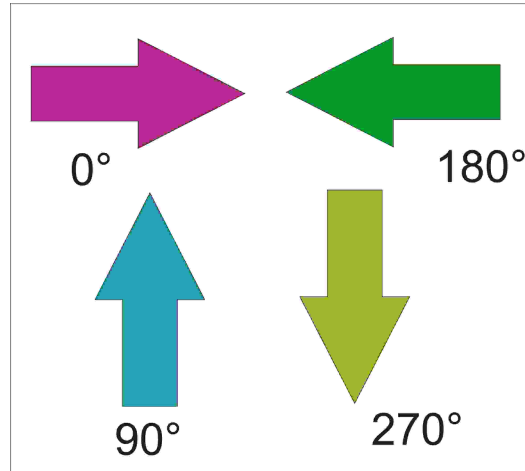


Abbildung 3.19: Rotationen der Formen

Die Reflektion (von  $0^\circ$  auf  $180^\circ$ ) ist eine der möglichen Rotationen. Für die Betrachtung der Rotationen eines dreidimensionalen Objekts eignen sich Kugelkoordinaten  $P(r, \phi, \theta)$ . Für die Koordinaten  $x, y, z$  gilt:

$$x = r \sin \theta \cos \phi \quad (3.15)$$

$$y = r \sin \theta \sin \phi \quad (3.16)$$

$$z = r \cos \theta \quad (3.17)$$

Die Cauchy-Schwarzsche Ungleichung gilt für ein abstraktes Skalarprodukt:

$$| \langle v, w \rangle |^2 \leq \langle v, v \rangle \cdot \langle w, w \rangle \quad (3.18)$$

Die Gleichung kann im Fall von  $x, y \neq 0$  umgeformt werden:

$$\frac{| \langle v, w \rangle |}{\sqrt{\langle v, v \rangle \cdot \langle w, w \rangle}} \leq 1 \quad (3.19)$$

$$\cos \phi = \frac{| \langle v, w \rangle |}{\sqrt{\langle v, v \rangle \cdot \langle w, w \rangle}} \quad (3.20)$$

Es gilt im Weiteren die Parallelogrammgleichung, wobei  $|x|^2 = \langle x, x \rangle$  das Skalarprodukt definiert.  $u$  bezeichnet die Drehung des Objekts.

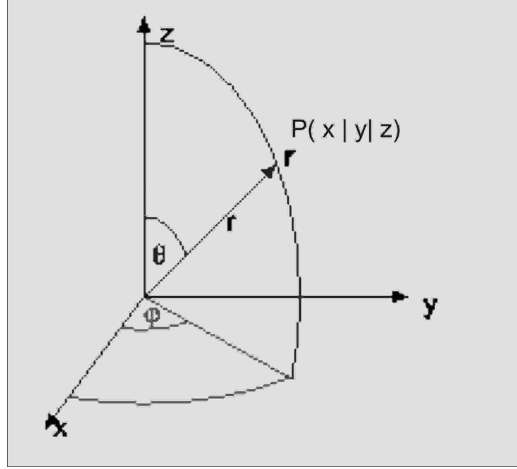


Abbildung 3.20: Kugelkoordinaten

$$\begin{aligned}
 |x_{\star} - \underbrace{u}_{\text{Drehung}} y_{\star}| &= \langle x - uy, x - uy \rangle = \langle x, x - uy \rangle - \langle uy, x - uy \rangle \\
 &= \langle x, x \rangle - \langle x, uy \rangle - \langle uy, x \rangle + \langle y, y \rangle \\
 &= \langle x, x \rangle - 2u \langle x, y \rangle + \langle y, y \rangle
 \end{aligned} \tag{3.21}$$

Die Koordinaten für eine Landmarke nach einer Rotation ergeben sich:

$$\begin{aligned}
 x'_1 &= x_1 \cos \theta - x_2 \sin \theta + 0x_3 \\
 x'_2 &= x_1 \sin \theta + x_2 \cos \theta + 0x_3 \\
 x'_3 &= 0x_1 + 0x_2 + 1x_3
 \end{aligned} \tag{3.22}$$

Es sind somit:

$$1 \quad \langle x, y \rangle = Rv \text{ mit } R \geq 0.$$

2 Die Drehmatrix  $u$

$$\|x - uy\| = \|x\|^2 + \|y\|^2 - \underbrace{(\langle xy \rangle u)}_{Rvu} + \underbrace{(\langle yx \rangle u)}_{Rvu} = \|x_{\star}\|^2 + \|y_{\star}\|^2 - 2Rvu. \tag{3.23}$$

Für  $u = v$  erreicht obige Gleichung das Minimum, für  $u = -v$  das Maximum. Das bedeutet, dass das Minimum dann vorliegt, wenn keine Rotation vorzunehmen ist, und das Maximum, wenn die beiden Objekte genau entgegengesetzt

---

zueinander liegen.

Zum Erreichen des Minimums ist  $y$  um den Winkel  $\phi$  und  $\theta$  zu drehen und um das Maximum zu erreichen um den Winkel  $\phi + \pi$  und  $\theta + \pi$ . Gesetz dem Fall  $R \geq 0$  kann Maximum und Minimum von keinem anderen  $u \in S$  erreicht werden. Durch die Eliminierung der Rotation wird die Distanz zwischen den Objekten minimiert.

Bei der Bestimmung der mittleren Form ist die Rotation der Objekte ( $n > 2$ ) auch dann zu beachten, wenn bei der Gruppe der zu untersuchenden Objekte keine Rotation zu eliminieren ist.

**Beispiel 14.** (2 Dreiecke) Seien die Dreiecke  $A(0|1), B(0|-2), C(3|-2)$  und  $D(0|3), E(0|-2), F(7|-2)$ . Die mittlere Form ist der Mittelwert  $M_1(0|2), M_2(0|-2), M_3(5|-2)$ .

Im Allgemeinfall kann man die mittlere Form nur bei zwei Objekten durch den Mittelwert der Landmarken bestimmen. Hat man mehr als zwei Objekte, so gestaltet sich die Bestimmung einer eindeutigen mittleren Form, die in optimalen Position zu allen Objekten liegt, aufwendiger [89].

**Behauptung . Für  $n > 2$  gilt Mittlere Form  $\neq$  Mittelwert.**

**Beispiel 15.** Für die drei Dreiecke

$$\begin{aligned}\mathfrak{L} &= \{\{1, 2\}, \{2, 3\}, \{1, 4\}\}, \\ \mathfrak{M} &= \{\{1, 0\}, \{-1, 2\}, \{-2, 1\}\}, \\ \mathfrak{N} &= \{\{4, -100\}, \{1, -1\}, \{-1, -2\}\}\end{aligned}$$

ergibt sich die mittlere Form

$$\hat{M} = \{\{1.33333, -34.\}, \{-0.666667, -0.666667\}, \{-1.33333, -1.66667\}\}.$$

Der Mittelwert hingegen ist  $M = \{\{2, -32.6667\}, \{0.6667, 1.3333\}, \{-0.6667, 1\}\}.$

**Beispiel 16.** Seien die drei dreidimensionale Objekte mit jeweils vier Punkten

$$\begin{aligned}\mathfrak{L} &= \{\{1, 2, 0\}, \{2, 4, 1\}, \{3, 5, 2\}, \{2, 3, 7\}\}, \\ \mathfrak{M} &= \{\{-1, 5, 0\}, \{2, 7, 1\}, \{3, -1, 2\}, \{2, -1, 8\}\}, \\ \mathfrak{N} &= \{\{-2, 6, -1\}, \{-3, 8, 1\}, \{-1, -1, 2\}, \{-200, -1, -1\}\}.\end{aligned}$$

Sie ergeben die mittlere Form

$$\hat{M} = \{\{-0.666667, -0.333333, -0.333333\}, \{-2.33333, -1., -0.333333\}, \{-2.33333, -1.66667, -0.666667\}, \{-68., -1., -5.33333\}\}.$$

Der Mittelwert hingegen ist

---


$$M = \{\{-0.666667, 4.333333, -0.333333\}, \{0.333333, 6.333333, 1\}, \\ \{1.666667, 1, 2\}, \{-65.333333, 0.333333, 4.666667\}\}.$$

Da es sich bei der mittleren Form nicht um eine einfache Berechnung des Mittelwertes handelt, benötigen wir einen entsprechenden Algorithmus zu ihrer Bestimmung.

### 3.5 Berechnung der mittleren Form

Man verwendet folgenden Algorithmus nach Ziezold(1994) [90] unter Anwendung des Skalarprodukts:

$$i = 1, \dots, n$$

$$\tilde{m} \mapsto w_i(\tilde{m}) = \begin{cases} \frac{\langle \tilde{m}, o_i \rangle}{|\langle \tilde{m}, o_i \rangle|} & \text{wenn } \langle \tilde{m}, o_i \rangle \neq 0 \\ 1 & \text{wenn } \langle \tilde{m}, o_i \rangle = 0 \end{cases} \quad (3.24)$$

$$\tilde{m} \mapsto T(\tilde{m}) = \frac{1}{n} \sum_{i=1}^n w_i(\tilde{m}) o_i \quad (3.25)$$

Rekursiv ergibt sich die Folge:

$$\tilde{m}_r = T(\tilde{m}_{r-1}), r = 1, 2, \dots \text{Iterationen.} \quad (3.26)$$

Abbruchkriterium des Algorithmus:

$$\tilde{m} = T(\tilde{m}) \quad (3.27)$$

Vorteil des Algorithmus ist, dass er in sehr kurzer Zeit, meist in weniger als drei Durchgängen, die mittlere Gestalt oder Form der Objekte bestimmt.

**Behauptung .** *Durch den beschriebenen Algorithmus wird erreicht, dass die so errechnete mittlere Form optimal zu allen  $n$  Objekten liegt.*

**Beweis .** *Sei  $y$  irgendeine zentrierte  $k$ -eckige Form und seien  $x'_i = w_i(y)x_i$  für  $i = 1, \dots, n$  die zentrierten Objekte. Sind die  $x'_i$  und  $y$  in optimaler Lage, so ist der Abstand unabhängig von Verschiebung und Drehung und die Objekte liegen nach dem Algorithmus optimal zur mittleren Form. Im Folgenden wird dies erstmalig bewiesen:*

Zu zeigen:  $\frac{1}{n}f(y) = \frac{1}{n}f(y') \Leftrightarrow y = y'$

---


$$\begin{aligned}
\frac{1}{n}f(y) &= \frac{1}{n} \sum_{i=1}^n \delta(y, x_i) = \frac{1}{n} \sum_{i=1}^n |y - x'_i|^2 \\
&= \frac{1}{n} \sum_{i=1}^n (|y|^2 + |x'_i|^2 - 2\langle y, x'_i \rangle) \\
&= \frac{1}{n} \sum_{i=1}^n (|y|^2 + |x'_i|^2 - 2\langle y, \frac{1}{n} \sum_{i=1}^n x'_i \rangle) \\
&= \frac{1}{n} \sum_{i=1}^n (|y|^2 + |x'_i|^2 - 2\langle y, y' \rangle) \tag{3.28} \\
&= \frac{1}{n} \sum_{i=1}^n (|y'|^2 + |x'_i|^2 - 2\langle y, y' \rangle \Leftrightarrow y = y') \\
&= \frac{1}{n} \sum_{i=1}^n |y' - x'_i|^2 \\
&= \frac{1}{n} f(y')
\end{aligned}$$

□

*Bemerkung .* Da in unseren Anwendungen Rotationen von Objekten nicht vorkommen, wird ab jetzt die Vorform nach Zentrierung und Normierung als die zu untersuchende Form genommen.

**Überlegungen .** Wenn man zwei disjunkte Mengen von Objekten  $\mathbb{M}$  und  $\mathbb{M}^a$  hat, die beide durch eine mittlere Form charakterisiert werden, so sollte ein Objekt  $o_i \in \mathbb{M}$  einen geringeren Abstand zu der mittleren Form oder Gestalt dieser Menge aufweisen als ein Objekt  $o_i^a \in \mathbb{M}^a$ . Ein noch nicht eingeordnetes Objekt  $o_i^b$  kann man also der Menge zuordnen, zu dessen mittleren Form es den geringsten Abstand besitzt.

Im folgenden Bild 3.21 ist die mittlere Form der Tumore von 24 Patienten der Pilotuntersuchung von Giebel [32], sowie der Tumor einer der Patienten in der Frontalansicht dargestellt.

Diese mittlere Form muß nicht ein Objekt der Menge sein. Es handelt sich lediglich um einen Erwartungswert. Das bedeutet, im Durchschnitt über eine entsprechende Menge an Objekten wird in dieser Gruppe diese Form erwartet. Dies ist eine Analogie zum Erwartungswert von Zufallsprozessen.

**Beispiel 17.** Die Dreiecke  $\mathfrak{p}$  mit den Punkten  $A(1|0), B(2|0)$  und  $C(1|7)$  und  $\mathfrak{p}$  mit den Punkten  $P(-1|0), Q(-4|0)$  und  $R(-1|7)$  haben als Mittelwert bzw. mittlere Form das Dreieck  $\mathfrak{m}$   $M_1(0|0), M_2(-1|0)$  und  $M_3(0|7)$ . Keines der beiden Ausgangsdreiecke  $\mathfrak{p}$  und  $\mathfrak{q}$  entspricht dem mittleren  $\mathfrak{m}$

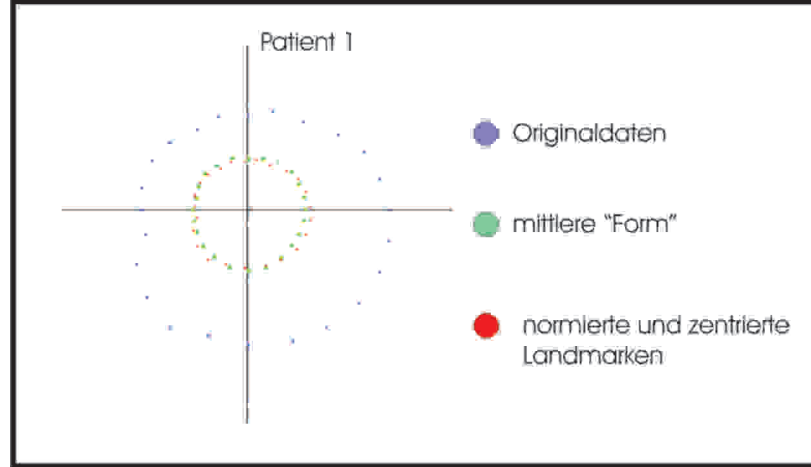


Abbildung 3.21: Patient Nr. 1 und die mittlere Form aus der Studie[32]

Ein anderer Ansatz, eine spezifische Form für eine Gruppe zu finden, ist der Median. Der Median ist eine der bereits in der Erhebung erfassten Formen.

Sei  $1_X(x), x \in \mathbb{R}^d$  die Indikatorfunktion [86]. Der Mittelwert, der Median und die Varianz der Indikatorfunktion führen zu einer deterministischen Funktion in  $\mathbb{R}^d$ .  $p_X(x) = E1_X(x) = P(1_X(x) = 1) = P(x \in X)$ .

**Definition 3.5.1.** Der Median ist definiert als  $Med_i(X) = \{m(x)\}_{x \in \mathbb{R}^m}$  mit  $m(x) = 1$ , wenn  $p_X(x) \geq \frac{1}{2}$  und anderenfalls  $m(x) = 0$ .

Die Wahrscheinlichkeit  $p_X(x)$  kann aus der Anordnung der Objekte errechnet werden [86]. Der Vorteil des Medians ist, dass es sich um ein Objekt handelt, das in der Realität beobachtet werden kann und gegenüber Ausreißern stabil ist.

**Beispiel 18.** Sei eine Menge von drei Objekten, ein Kreis, ein Fünfeck und ein Rechteck. Um den Median zu bestimmen, sind die Objekte in eine aufsteigende Anordnung zu überführen. Die unten abgebildete Anordnung spiegelt den steigenden Flächeninhalt wieder. Analog kann dies für den Umfang etc. durchgeführt werden. Der Median im Bild 3.22 ist dann das gelbe Rechteck.

## 3.6 Varianz

Die Varianz wird in der Formanalyse bisher eindimensional gefasst. Bei fehlender Kenntnis der Verteilung kann die Varianz zur Abschätzung der Wahrscheinlichkeit für eine Abweichung mittels der Tschebyscheffschen Ungleichung verwendet werden: Für irgendein  $t > 0$  :



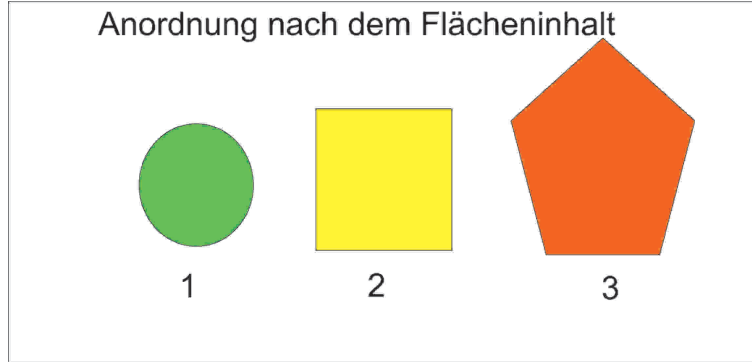


Abbildung 3.22: Bestimmung des Medians innerhalb einer Gruppe von Objekten

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2} \quad (3.29)$$

Die Varianz kann dazu dienen, eine zulässige Abweichung von der mittleren Form zu bestimmen. Damit kann der Anwender den zufälligen Fehler bei der Messung der Daten berücksichtigen.

**Beispiel 19.** In meiner ersten Untersuchung [32] bei Nierentumoren im Kindesalter habe ich 24 Tumore für die zweidimensionale Auswertung zur Verfügung gehabt und ausgewertet. Der durchschnittliche Abstand der 18 normierten und zentrierten Wilms-Tumore (Nephroblastome) beträgt 0,1304. Die Varianz beträgt 0,002. Die Abweichung von der mittleren Form ist nur einseitig zu betrachten, denn je näher ein Objekt an der mittleren Form der Wilms-Tumore liegt, desto eher ist von einem Wilms-Tumor auszugehen. Entsprechend wäre die Wahrscheinlichkeit bei Annahme einer symmetrischen Verteilung nur für einen größeren Wert als den Mittelwert ( $X > 0,1304$ ) zu berechnen:

$$P((X - 0.1304) \geq \frac{t}{2}) \leq \frac{4 \cdot 0,002^2}{t^2} .$$

### 3.7 Mittlere Varianz

Die Varianz ist eine reelle Zahl. Sie ist ein Maß für die gesamte Streuung. Bei mehrdimensionalen Daten ist es von Interesse, die Abweichung von der mittleren Form je nach Landmarke und Dimension zu untersuchen. Analog zur "mittleren Form" kann die "mittlere Varianz" als der Erwartungswert der Distanz zur "mittleren Form" berechnet werden. Die "mittlere Varianz" ist dann die kleinste Distanz zwischen allen Objekten und der "mittleren Form".

---

**Definition 3.7.1. Mittlere Varianz:**  $\tilde{\sigma}$  wird als Erwartungswert der Abstände von der mittleren Form bezeichnet, also als Mittel von  $m - x_1 \dots m - x_n$ , wenn:

$$\sum_{i=1}^n d(m - x_i, \tilde{\sigma})^2 = \inf_{\alpha \in \Xi} \sum_{i=1}^n d(m - x_i, \alpha)^2, \forall \alpha \in \Xi \quad (3.30)$$

Das bedeutet, dass die "mittlere Varianz" definiert ist als der Mittelwert der Distanzen zwischen der mittleren Form und den Objekten.

### 3.8 Alternative Berechnung einer mittleren Form mittels neuronaler Netze

Die mittlere Form kann auch mit Hilfe eines einschichtigen neuronalen Netzes berechnet werden, welches die einzelnen Landmarken der Objekte gewichtet und die Varianz minimiert. Vorteil dieses Vorgehens ist es, nicht bei jeder Dimension der Landmarken und jeder Landmarke von einer gleichen Transformation auszugehen.

Die mittlere Form in den bisherigen Anwendungen wie Muscheln [90] ging von einer vergleichbaren Transformation aller Landmarken in allen Dimensionen aus. Nicht nur die Objekte werden im Gegensatz zu dem Algorithmus [90] unterschiedlich gewichtet, sondern jede Landmarke in jeder Dimension: Eine vorhergehende Normierung und Zentrierung ist bei diesem Vorgehen nicht unbedingt notwendig, da die Objekte entsprechend der Objektgruppe optimal während des Verfahrens angepasst werden. Zu vergleichende Objekte sind entsprechend der Gewichtungen dann zu transformieren.

Zusätzlich zur Minimierung der Varianz wird der  $u$ -Wert berechnet. Es wird die Form gewählt, die eine optimale Differenzierung zu der entsprechend anderen Objektgruppe zulässt.

Die Minimierung der Varianz führt dazu, dass mit geringer werdender Varianz der stochastische Prozess abbricht. Daher ist bei der Minimierung der Varianz im allgemeinen eine Verringerung der Differenzierbarkeit hinsichtlich anderer Objektgruppen zu erwarten. Mittels der Formel von Tschebyscheff ergibt sich  $P(|X - m| \geq \epsilon) = 0, \forall \epsilon > 0$ . Für  $P(|x - m| > 0) = P(x \neq m) = 0$ , für das Gegenereignis  $P(x = m) = 1$ . Das bedeutet, dass nur noch ein Ereignis nämlich  $A = \{x = m\}$  eintreten kann. Bei einer diskreten Verteilung bedeutet das, dass alle beobachteten Objekte dem Objekt  $m$  mit 100%iger Wahrscheinlichkeit entsprechen.

Ein solches Vorgehen wird bei der Formerkennung angewandt, wenn die zu suchende Form bereits bekannt ist und lediglich in einer Gruppe von Formen wiederzuerkennen ist. Minimierung in einer der beiden Gruppen führt damit

---

nicht immer zu einer optimalen Differenzierung der Objektgruppen.

---

## Kapitel 4

# Differenzierung und Relevanz der Landmarken

Die bisher als Formanalyse bezeichneten Ansätze beschäftigen sich größtenteils mit Formen, die dem Anwender bereits bekannt und vertraut sind. Bei der Bildverarbeitung geht es meist darum, das Bildmaterial für die Erkennung der Formen auf das Wesentliche zu komprimieren oder darum im Bild, eine Form wieder zu finden.

Artikel über die statistische Formanalyse sind meist theoretischer Natur. Trotz der durchaus vorhandenen Anwendungsmöglichkeiten fehlt es an entsprechenden Umsetzungen.

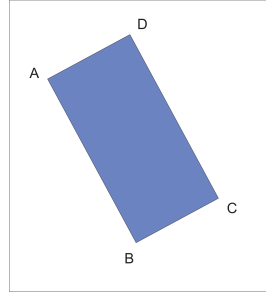
Im Bereich der Radiologie ist das besonders erstaunlich, da gerade in diesem Bereich ein statistischer Ansatz für die Auswertung von Formen nahe liegend ist. Zum einen handelt es sich um Formen, die einem stochastischen Prozess entstammen, und im Gegensatz zu Organen oder Knochen keine klare Form besitzen, und zum anderen ist es möglich die Informationen für eine Auswertung direkt aus dem vorhandenen Bildmaterial zu entnehmen. Oft können vom Mediziner a priori keine oder nur unzureichend relevante Parameter angegeben werden. In diesem Fall können auch vorgegebene Formen wie die Kugel oder das Ellipsoid benutzt werden, um regelmäßig verteilte Landmarken zu definieren. Außerdem müssen körperlich gegebene Beschränkungen der Form beachtet werden. Nierentumore werden in ihrem Wachstum zum Beispiel durch die Wirbelsäule und die umliegenden Organe eingeschränkt. Dies wirkt sich auch auf ihre mittlere Form aus. Man könnte also annehmen, dass einige der Landmarken für die Bestimmung der mittleren Form wichtiger sind als andere.

### 4.1 Wiederherstellung von Formen

Um eine bekannte Form wiederherzustellen, gibt es zum einen die Möglichkeit, die vorhandenen Landmarken zu verwenden, um die übrigen Landmarken zu ermitteln. Das heißt, es wird eine Zuordnung  $f$  ermittelt, die ausgehend von

den Landmarken  $x_i$  für  $i = 1 \dots m$  die Landmarken  $x_i$  für  $i = m + 1 \dots k$  bei  $k$  Landmarken bestimmt.

**Beispiel 20.** Bei einem Rechteck  $:A(1|7), B(2|1), C(3|4)$  ist  $D$  zu bestimmen. An  $A$  ist der Vektor  $v$  der Strecke von  $B$  nach  $C$  anzulegen. Demnach ergibt sich  $D$  aus den Koordinaten von  $A$  um den Vektor  $v = (1, 2)$  verschoben:  $D(2|9)$ .



Beispiel für die Wiederherstellung der Form bei einem Rechteck

Der andere Ansatz zur Wiederherstellung einer Form bedient sich eine Funktion  $F$ , die die Form im  $R^3$  beschreibt. Sämtliche Landmarken sind in beliebigen Abständen über diese Funktion  $F$  zu bestimmen.

**Beispiel 21.** Anhand der 16 Raubmörder ist aus der Entfernung von der ersten Tat zur zweiten Tat  $x_1$  die Entfernung von der zweiten zu dritten Tat  $y$  zu bestimmen. Es wird die Lineare Regression angesetzt:  $y = \beta_0 + \beta_1 x_1$

In vielen Bereichen, wie bereits bei den Raubmördern, ist nicht von einer deterministischen Form auszugehen. Die meisten beobachteten Formen entstammen einem stochastischen Prozess.

## 4.2 Test auf Differenzierbarkeit

Um die "mittlere Form" oder "mittlere Varianz" als ein verwendbares Charakteristikum einer Gruppe verwenden zu können, ist zu prüfen, ob diese eine Differenzierbarkeit von den jeweils anderen Gruppen ermöglicht.

Beim Testen der Annahme auf Zufälligkeit gehen wir einen Fehler ein, der mit  $\alpha$  beschrieben wird. In  $\alpha \cdot 100$  Fällen wird die Zufälligkeit zwar stimmen, aber wir verwerfen die Annahme der Zufälligkeit fälschlicherweise. Es wird vom Fehler 1. Art gesprochen.

Einen Fehler 2. Art gehen wir ein, wenn wir die Anfangsannahme fälschlicherweise beibehalten. Die Wahrscheinlichkeit hierfür wird mit  $\beta$  beschrieben. Sie ist aufgrund des Vorgehens bei Tests meist nicht berechenbar. Es sei denn, man nimmt einen weiteren Test mit einer anderen konkurrierenden Anfangsannahme.

### 4.2.1 Test der mittleren Form

Der Test von Ziezold [90] kann in folgenden Schritten beschrieben werden:

---

### 1. Schritt: Definition der Untermengen

Es gibt eine Menge  $M = \{o_1, \dots, o_N\}$  :

Die Untermenge  $A$  ist eine Realisierung der Verteilung  $P$  und die Untermenge  $B$  ist eine unabhängige Realisierung der Verteilung  $Q$ .

*Hypothese:*  $H_0 : P = Q$

*Alternative:*  $H_1 : P \neq Q$

Setze ein *Signifikanzniveau*  $\alpha$ . Wenn die Wahrscheinlichkeit für  $H_0$  niedriger als das Signifikanzniveau ist, wird  $H_0$  abgelehnt und  $H_1$  angenommen.

### 2. Schritt: Berechnen der "mittleren Form"

Die "mittlere Form" wird berechnet durch den Algorithmus von Ziezold (1994)[90]. Im Weiteren bezeichnet  $m_0$  die "mittlere Form" der Untermenge  $A$ .

### 3. Schritt: Berechnen des $u$ -Wertes

$$u_0 = \sum_{j=1}^n \text{card}(b_k : d(b_k, m_0) < d(a_j, m_0))$$

**4. Schritt: Bestimmung aller Möglichkeiten die gesamte Menge aller Objekte in zwei Menge im gleichen Verhältnis aufzuteilen**

**5. Schritt: Vergleichen des  $u_0$ -Wertes mit allen sich aus den möglichen Aufteilungen ergebenden  $u$ -Werten. Berechne den Rang (ein niedriger  $u$ -Wert bedeutet ein niedriger Rang).**

**6. Schritt: Berechne den  $p$ -Wert für  $H_0$**

$p_{r=i} = \frac{r}{\binom{N}{n}}$  for  $i = 1, \dots, \binom{N}{n}$ , wobei  $r$  der Rang ist, für welchen wir Gleichverteilung annehmen.

Diesen Rechenvorgang müssen wir bei diesem Test auch für die komplementäre Menge  $B$  durchführen, da ansonsten ein Verwerfen der Hypothese  $H_0$  lediglich eine Folge der unterschiedlichen Varianz innerhalb der beiden Mengen sein kann. Vorteil des Tests ist darin zu sehen, dass zum Prüfen des  $u$ -Wertes keine Verteilung angenommen wird, die es weiter zu prüfen gilt, sondern lediglich die

---

Gleichverteilung von zufälligen  $u$ -Werten.



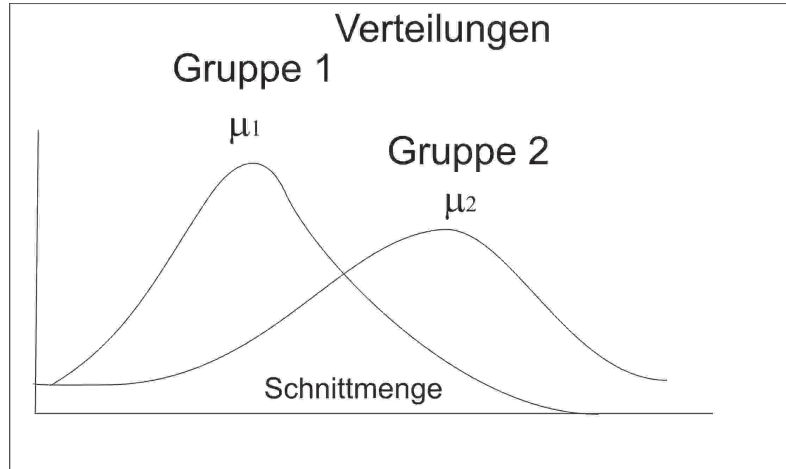


Abbildung 4.1: Darstellung zweier Gruppen mit unterschiedlicher Varianz

Nachteilig erweist sich bei der Anwendung des Tests bei meiner ersten Untersuchung bei zweidimensionalen Tumoren, dass die "mittlere Form" des jeweiligen Komplements zu berechnen und in der jeweiligen Disziplin zu interpretieren ist. Die Restgruppe ergibt sich in der Anwendung häufig als zufälliges Ergebnis der Erhebung und kann in Folge ihrer Heterogenität nicht als eine Gruppe in der Disziplin gedeutet werden. Dieser Nachteil des Verfahrens, einer Berechnung mit Referenz zu beiden Gruppen, lässt sich durch eine entsprechende Prüfung der Varianz beheben.

#### 4.2.2 Test der Varianz

Der Test einer mehrdimensionalen Varianz kann durch folgendes Verfahren erstmalig durchgeführt werden:

##### 1.Schritt: Definition der Untermengen

Es gibt eine Menge  $M = \{o_1, \dots, o_N\}$  :

Die Untermenge  $A$  ist eine Realisierung der Verteilung  $P$  und die Untermenge  $B$  ist eine unabhängige Realisierung der Verteilung  $Q$ . Die -Untermenge  $A$  hat die Varianz  $\sigma_1^2$  und  $B$  die Varianz  $\sigma_2^2$ .

$$\text{Hypothesis:} \quad H_0 : \sigma_1^2 = \sigma_2^2$$

$$\text{Alternative:} \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

---

Setze ein *Signifikanzniveau*  $\alpha$ . Wenn die Wahrscheinlichkeit für  $H_0$  niedriger als das Signifikanzniveau ist, wird  $H_0$  abgelehnt und  $H_1$  angenommen.

## 2. Schritt: Berechnung der Varianz

Die mehrdimensionale Varianz wird erstmalig in dieser Untersuchung analog zur "mittleren Form" in [90] berechnet. Sei  $\sigma_1^2$  die mehrdimensionale Varianz der Untermenge  $A$ .  $\sigma_2^2$  ist dann die berechnete mehrdimensionale Varianz von  $B$ .

## 3. Schritt: Berechnen des $F$ -Wertes

$$F = \frac{|\hat{\sigma}_1^2|}{|\hat{\sigma}_2^2|}$$

**4. Schritt: Berechnung aller oder eines zufälligen Teils von möglichen Aufteilung der Gesamtgruppe mit der gleichen Proportion von  $A$  und  $B$  zueinander.**

**5. Schritt: Vergleiche den  $F_0$ -Wert mit allen weiteren in Schritt 4 berechneten  $F$ -Werten. Berechne den Rang (ein geringer  $F$ -Wert bedeutet ein geringer Rang).**

## 6. Schritt: Berechne den $p$ -Wert für $H_0$

$p_{r=i} = 1 - \frac{r}{\binom{N}{n}}$  for  $i = 1, \dots, \binom{N}{n}$ , wobei  $r$  der Rang ist für welchen wir eine Gleichverteilung annehmen und sich eine Wahrscheinlichkeit mit  $p_{r=i} = \frac{1}{\binom{N}{n}}$  ergibt.

Dieser Test ist lediglich als eine Ergänzung des Tests von Ziezold(1994) [90] zu sehen. Der Test ist lediglich in der Lage die Varianzunterschiede der beiden Gruppen zu prüfen.

### 4.2.3 Test der mittleren Form und der Varianz

Der Test von Ziezold(1994)[90] weist den Nachteil auf, dass er sowohl hinsichtlich der mittleren Form der Menge  $A$ , als auch hinsichtlich der mittleren Form der komplementären Menge  $B$  anzuwenden ist, um eine Aussage über die Verwendbarkeit der mittleren Form erhalten zu können. Im vorherigen Test wurde zusätzlich zu dem Test der mittleren Form die Prüfung der Varianz vorgestellt. Im nun folgenden Test wird die mittlere Form unter Beachtung der Varianz in einem Vorgehen geprüft:

#### 1.Schritt: Definition der Untermengen

Die Untermenge  $A$  ist eine Realisierung der Verteilung  $P$  und die Untermen-

---

ge  $B$  ist eine unabhängige Realisierung der Verteilung  $Q$ .

*Hypothese:*  $H_0 : P = Q$

*Alternative:*  $H_1 : P \neq Q$

Setze ein *Signifikanzniveau*  $\alpha$ . Wenn die Wahrscheinlichkeit für  $H_0$  niedriger als das Signifikanzniveau ist, wird  $H_0$  abgelehnt und  $H_1$  angenommen.

## 2. Schritt: Berechnen der "mittleren Form"

Die "mittlere Form" wird berechnet durch den Algorithmus von Ziezold (1994)[90]. Im Weiteren bezeichne  $m_0$  die "mittlere Form" der Untermenge  $A$ .

## 3. Schritt: Berechnen des $u$ -Wertes

Im Vergleich zum Test von Ziezold(1994) wird die Varianz in Form der Standardabweichungen in der Menge  $A$  und  $B$  berücksichtigt.

$$u_0 = \sum_{j=1}^n \text{card}(b_k : d(b_k, m_0) + \sigma_B < d(a_j, m_0) + \sigma_A)$$

## 4. Schritt: Bestimmung aller Möglichkeiten die gesamte Menge in zwei Menge im gleichen Verhältnis aufzuteilen

**5. Schritt: Vergleichen des  $u_0$ -Wertes mit allen sich aus den möglichen Aufteilungen ergebenden  $u$ -Werten. Berechne den Rang (ein niedriger  $u$ -Wert bedeutet ein niedriger Rang).**

## 6. Schritt: Berechne den $p$ -Wert für $H_0$

$p_{r=i} = \frac{r}{\binom{N}{n}}$  for  $i = 1, \dots, \binom{N}{n}$ , wobei  $r$  der Rang ist, für welchen wir Gleichverteilung annehmen.

Vorteil des Tests ist darin zu sehen, dass zum Prüfen der Aussage nicht die mittlere Form der komplementären Menge  $B$  berechnet werden muss. Die Berücksichtigung der Varianz in Form der Standardabweichung innerhalb des Test kann sich bereits theoretisch ergeben. So ist davon auszugehen, dass Tumore, die nur in der Nähe eines Organs wachsen, eine geringere Varianz haben als Tumore, die überall im menschlichen Körper wachsen können.

## 4.2.4 Entscheidungsbäume

In einer Untersuchung des Maßregelvollzugs Hadamar, der sich speziell der Behandlung von drogen- und alkoholabhängigen Straftätern widmet, wurden aus

den Entlassungsjahrgängen 1998 bis 2003 144 Patienten erfasst. Für 107 Patienten konnte ein vergleichbarer Zeitraum nach Entlassung aus dem Maßregelvollzug sicher gestellt werden[49].

Für die statistische Ermittlung von Risiko- und Schutzfaktoren, sowie der entsprechenden Klientel wurden die Entscheidungsbäume verwendet. Entscheidungsbäume ermitteln über ein Testverfahren wie den  $\chi^2$ -Test das statistisch in der jeweiligen Gruppe relevanteste erklärende Merkmal. Die Daten werden anhand der Ausprägungen des erklärenden Merkmals aufgeteilt und wieder in den sich ergebenden Untergruppen nach weiteren gemäß dem  $\chi^2$ -Test relevanten Merkmalen untersucht. Es handelt sich um ein schrittweise bivariates Verfahren. Es werden nicht wie bei multivariaten Verfahren alle Merkmale gleichzeitig im Verfahren verarbeitet. Der  $\chi^2$ -Test setzt voraus, dass die Merkmalsausprägungen in sinnvolle Kategorien zusammengeführt werden können. Die Kategorien sollten mindestens eine zu erwartende Fallzahl von fünf nicht unterschreiten.

Anstatt des  $\chi^2$ -Test kann auch ein anderes beliebiges Testverfahren wie das von Ziezold (1994) [90] beschriebene verwendet werden. Durch die in dieser Arbeit erstmalige Kombination der Entscheidungsbäume mit dem Test von Ziezold (1994) [90] wird die Betrachtung auf Subgruppe innerhalb der differenzierbaren Gruppen gelenkt. Die Aufteilung erfolgt dann zunächst anhand der Gruppe mit der "mittleren Form", zu der die meisten gehören und die signifikant zur Unterscheidung von den übrigen nicht zu der Gruppe zugehörigen beiträgt. Dieser Schritt erfolgt dann wieder für die jeweiligen Untergruppen von Objekten.

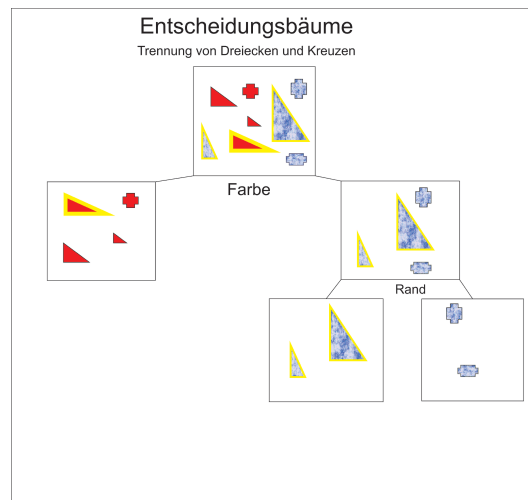


Abbildung 4.2: Zuordnung zur Diagnose anhand der Entscheidungsbäume

Der bisherige Ansatz des  $\chi^2$ -Tests wurde bereits in der seit 2006 begonnenen Untersuchung des Maßregelvollzugs Hadamar [25] angewandt und vorgestellt. Folgende Auswertung der Entscheidungsbäume ist ein anschauliches Beispiel für den schrittweise angewandten  $\chi^2$ -Test.

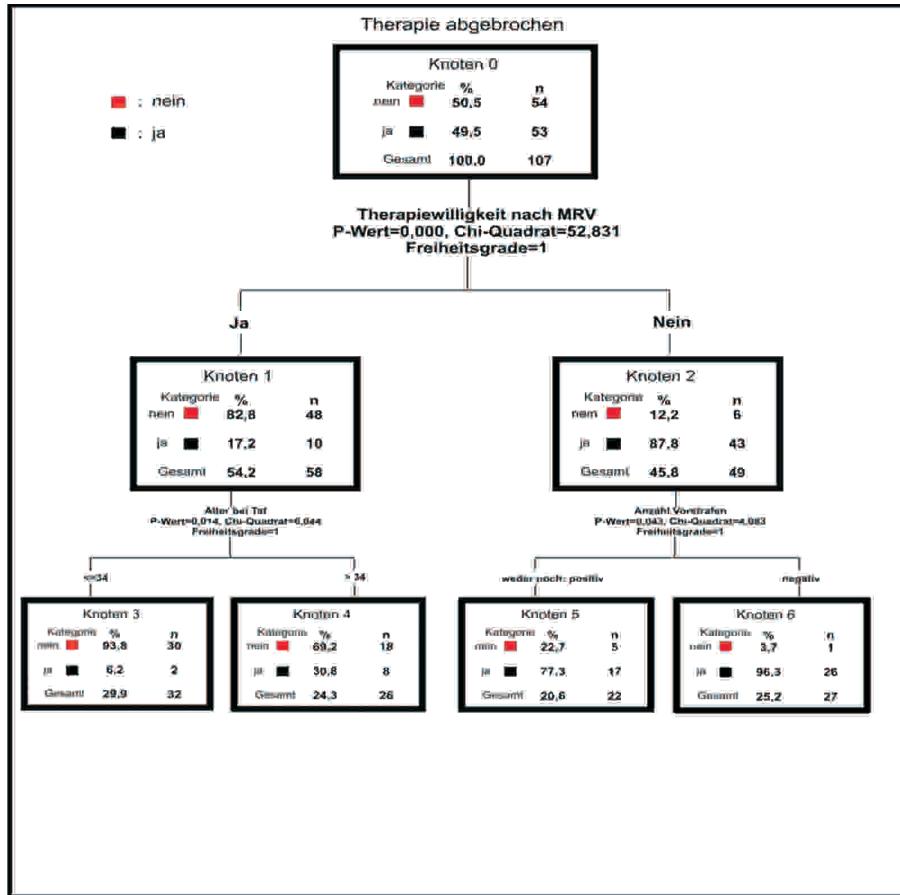


Abbildung 4.3: Erklärung des Therapieabbruch im Maßregelvollzug Hadamar anhand der Entscheidungsbäume

In der Gruppe von 107 Patienten mit einem Mindestzeitraum von zwei Jahren nach Entlassung aus dem Maßregelvollzug stellt sich als entscheidender Faktor die Therapiewilligkeit heraus. In der Gruppe der Therapiewilligen ergibt sich das Alter unter/gleich und über 34 Jahre als weiterer Faktor. Die jüngeren Patienten sind ebenfalls eher diejenigen, die die Therapie durchhalten. Die Therapieunwilligen unterscheiden sich durch das Verhalten im Maßregelvollzug. Die, die negativ im Maßregelvollzug auffallen, brechen zu 96,3% und die, die neutral oder positiv auffallen, zu 77,3% die Therapie ab. Das Ergebnis hinsichtlich der Therapiewilligkeit ist bei beiden Auswertungen den Abläufen

---

des Vollzugs geschuldet. Patienten, die als therapieunwillig eingestuft werden, werden natürlich eher die Therapie abbrechen. Der Maßregelvollzug wird von selbst eher den Abbruch anregen. Die geringere Anzahl an Therapieabbrechern unter den als therapiewillig eingestuften Patienten zeigt zumindest, dass diese Patienten nicht von selbst dann die Therapie abbrechen. Die bei der Gruppe der therapiewilligen Patienten vorgenommene Unterteilung in diejenigen, die jünger als/oder gleich 33 bzw. 34 Jahre alt sind, und diejenigen, die älter sind, zeigt, dass für die erfolgreiche Beendigung der Therapie eher die jüngeren zu erreichen sind als die älteren. Der Grund ist darin zu sehen, dass bei den jüngeren noch eine gewisse Entwicklung zu erwarten ist und mit der Mitte des dreißigsten Lebensjahrs noch ein neuer Lebensabschnitt mit Partnerschaft und beruflicher Integration beginnen kann.

Bei den Therapieunwilligen zeigt sich, dass die Wahrscheinlichkeit für einen Therapieabbruch mit der strafrechtlichen Vorbelastung wächst. Je stärker jemand strafrechtlich vorbelastet ist, desto weniger kann er zur Therapie bewegt werden und damit zu einem Abbruch seiner bisherigen kriminellen Karriere.

## 4.3 Landmarkentest

Generell stellt sich nicht nur die Frage nach der Differenzierbarkeit der Objekte, sondern die nach der Relevanz bestimmter Landmarken. Für den Mediziner könnte die Relevanz bestimmter Landmarken dazu führen, dass bestimmte Regionen des Tumors stärker bei der Diagnose zu beachten sind, für den Ingenieur einige Sensoren bei der Bestimmung der Geruchsbelastung und Geruchssubstanz je nach Anwendung eine unterschiedliche Bedeutung besitzen und für den Kriminologen und Kriminalisten einige Orte des Tathergangs.

### 4.3.1 Test auf Einfluss der Landmarken auf die Form

Im Testverfahren [91] unterscheidet Ziezold die Unabhängigkeit der Landmarken von den anderen Landmarken hinsichtlich der Form innerhalb einer bestehenden Gruppe.

Da es sich nicht um die stochastische Unabhängigkeit handelt, wird hier von dem Einfluss der Landmarken auf die Form gesprochen.

Folgende Hypothesen werden aufgestellt:

$H_0$ : Die  $k$ -te der  $m$  Landmarken wird hinsichtlich der "Formen" nicht von den anderen Landmarken beeinflusst.

$H_1$ : Die  $k$ -te der  $m$  Landmarken wird hinsichtlich der "Formen" von den anderen Landmarken beeinflusst.

---

Die Anfangsannahme  $H_0$  wird verworfen, wenn mit einem großen Abstand der Formen ein großer Abstand der herausgenommenen  $k$ -ten Landmarke einhergeht.

Es wird ein geeignetes  $p$ -Quantil für die Menge der Distanzen zwischen den  $n$  Formen gewählt. Die Herausnahme der  $k$ -ten Landmarke wird mit einem Apophroph gekennzeichnet:

$$M = \{d(o'_i, o'_j) : 1 \leq i < j \leq n\}.$$

Mit  $c_p$  sei das  $p$ -Quantil bezeichnet und mit  $n_p \approx p \cdot \binom{n}{2}$  ist die Anzahl an Distanzen bei  $i < j$  in dem jeweiligen Quantil zu verstehen.

Entfernt man die erste Landmarke, so ergibt sich:

$$A_p^1 = \frac{1}{n - n_p} \sum_{\{d(o'_i, o'_j) > c_p\}} \frac{|l_k^i - l_k^j|}{d(o'_i, o'_j)}.$$

Je größer der Wert von  $A_p^1$  ist, desto eher ist von einem Verwerfen der Nullhypothese auszugehen. Für die Prüfung werden wieder alle  $N$  Möglichkeiten berechnet, eine Landmarke aus der Menge der Landmarken zu ziehen, und der Rang  $R_p$  von  $A_p^1$  in der so berechneten Menge  $\{A_p^1 \dots A_p^N\}$  bestimmt.

Unter der Annahme, dass der so bestimmte zufällige Rang gleich verteilt ist, wird  $H_0$  bei einem Signifikanzniveau  $\alpha$  verworfen, wenn:

$$\pi_p = \frac{N - (R_p - 1)}{N} \leq \alpha.$$

Wird  $H_0$  verworfen, wird mit der Alternative  $H_1$  davon ausgegangen, dass die entsprechende Landmarke von den übrigen Landmarken abhängig ist. Mit anderen Worten von den übrigen Landmarken beeinflusst wird.

**Beispiel 22.** In der Diplomarbeit [32] wurde das Testverfahren auf 18 zweidimensional erfasste Wilms-Tumore angewendet. Bei allen 24 Landmarken war die Hypothese einer Nicht-Beeinflussung durch die übrigen Landmarken bei einem Quantil von 65% und einer Wahrscheinlichkeit  $p = 0,30$  aufrecht zu erhalten[23]. Eine weitere Anwendung des Tests beim explorativen Ansatz der Erhebung ist nicht sinnvoll. Den einzelnen Landmarken kann ohnehin keine medizinische Bedeutung bei der Beschreibung nur einer Gruppe beigemessen werden.

Dieser Test ist nur in der Lage innerhalb einer vorhandenen Gruppe Aussagen über die Landmarken zu treffen. In allen drei Anwendungen benötigen wir

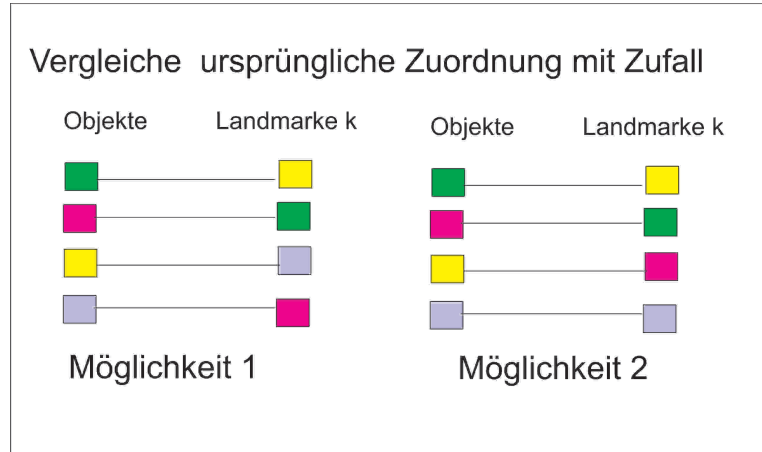


Abbildung 4.4: Vorgehen des Tests

die Differenzierung zu den jeweilig anderen Gruppen. Darüber hinaus ist die Wahl der Landmarken in allen hier untersuchten Anwendungsgebieten rein explorativ. Nur im Falle von theoretischen Vorüberlegungen erscheint dieser Test geeignet, insbesondere wenn anatomisch relevante Landmarken zur Verfügung stehen wie in der exemplarischen Untersuchung von Ziezold (2003) [91].

### 4.3.2 Selektion einer Auswahl an Landmarken

Der bisherige Test von Ziezold (2003) [91] versucht die Abhängigkeit der Landmarken innerhalb der Gruppen zu untersuchen. Wesentlich bedeutsamer ist die Abhängigkeit der Landmarken zur Differenzierung der Gruppen. Erst in der Pilotuntersuchung von Giebel [32] wurden solche Ansätze vorgestellt.

Um Landmarken für die Unterscheidung auszuwählen, kann rein explorativ folgendes Vorgehen durchgeführt werden:

#### *Landmarkenauswahl*

**1.Schritt:** Man wählt die Anzahl der Landmarken  $k$  für die Untersuchung.  $k$  ist dabei so zu wählen, dass eine Form des zu untersuchenden Objektes erkennbar ist, zumindest  $k > 2$ .

**2.Schritt:** Man geht alle Möglichkeiten  $k$  aus  $n$  Landmarken zu ziehen oder nur eine zufällige Anzahl davon durch

**3.Schritt:** Man bestimme für jede Konfiguration aus  $k$ -Landmarken den  $u_0$ -Wert für die Unterscheidung der zu untersuchenden Teilgruppe bei einer Gesamterhebung von  $N$



---

Um die Brauchbarkeit einer Landmarkenauswahl zu prüfen, ist es sinnvoll den bestehenden Datensatz aufzuteilen und die Aussagen aus dem bekannten Datensatz auf den jeweils unbekannten Datensatz zu übertragen. Bei der Übertragbarkeit des aus den bekannten Daten gebildeten Modells auf unbekannte Daten wird von der Vorhersage gesprochen.

*Bestimmung einer für die Vorhersage geeigneten Konfiguration*

**1.Schritt:** Der Datensatz wird in zwei zufällige Teile aufgeteilt. Der erste Part dient zur Ermittlung einer Erklärung in Form einer Zuordnung, der zweite Part dient zur Vorhersage, damit Anwenden der Zuordnung auf noch unbekannte Daten. Je nach der realen Situation kann der Datensatz aufgeteilt werden: Die Simulation der Vorhersage durch das Teilen der Daten kann der jeweiligen Realität angepasst werden.

**2.Schritt:** Man nimmt die Konfiguration im ersten Datensatz mit dem geringsten  $u_0$ -Wert. Dieser wird zum einem mit Referenz der mittleren Form der Gruppe  $A$  und zum anderen mit Referenz der mittleren Form der Gruppe  $B$  berechnet.

**3. Schritt:** Man betrachtet den  $u_0$ -Wert dieser Konfiguration. Eine verwendbare Konfiguration sollte auch bei dem zweiten noch nicht verwendeten Datensatz einen niedrigen  $u_0$ -Wert aufweisen.

**Beispiel 23.** Es werden folglich in meiner ersten Pilotuntersuchung [32] für die Unterscheidung der Nephroblastome (Wilms-Tumore) zu den Neuroblastomen im **1.Schritt** die 18 Nephroblastome in zwei Gruppen zu je neun zufällig zugeordnet. Aufgrund der geringen Menge an Neuroblastomen werden diese sowohl zur Erstellung des Modells als auch für die Prognose verwendet.

Es werden im **2.Schritt** anhand der ersten Gruppe die Fünferkonfigurationen gewählt, die den niedrigsten  $u_0$ -Wert in der Unterscheidung Nephro- zu Neuroblastomen haben. Von diesen ausgewählten Konfigurationen wird dann diejenige gewählt, die den kleinsten  $u_0$ -Wert in der Unterscheidung Neuroblastom zu Nephroblastom hat. Insgesamt tritt  $u_0 = 0$  zu 888 mal bei Nephro- zu Neuroblastomen auf.

$$M_{5er-Konfiguration} = \{(0, 1, 3, 13, 18), \dots (16, 17, 18, 19, 22)\}$$

$$|M| = 888$$

Von diesen werden diejenigen oder nur die eine gewählt mit dem niedrigsten  $u_0$ -Wert in der Richtung Neuroblastom zu Nephroblastom. In der vorliegenden Konstellation ergibt sich nur die Landmarkenkonfiguration  $L_{Auswahl} = \{3, 14, 15, 19, 22\}$  mit  $u_0 = 5$  für Neuroblastom zu Nephroblastom. Die gleiche

---

Konfiguration wird nun in der zweiten Stichprobe innerhalb des **3.Schrittes** betrachtet. Es ist festzustellen, dass der  $u_0$ -Wert für Nephro- zu dem Neuroblastom 4 beträgt und von Neuroblastom zu Nephroblastom 5. Bei derart niedrigen  $u_0$ -Werten kann von der Verwendbarkeit der Konfiguration gesprochen werden.

Zu beachten ist bei der Verwendbarkeit des Resultats, dass aufgrund der geringen Fallzahl von Neuroblastomen diese zufällige Trennung nicht für die Neuroblastomen analog zu Nephroblastomen (Wilms-Tumore) durchgeführt werden konnte, so dass im zweiten Datensatz zum Prüfen der Landmarkenkonfiguration nur dieselben Neuroblastome zur Verfügung stehen. Die Ergebnisse sind in [38] dargestellt.

### 4.3.3 Vorzeichenrangtest

Die Abweichung der Landmarken von der jeweiligen mittleren Form können ebenfalls für eine Aussage zur Differenzierung der Gruppen verwendet werden. Vorteil des Vorgehens ist, dass die unterschiedliche Anzahl an Objekten in den Gruppen unbeachtet bleibt. Vorhanden sind Objekte mit Eigenschaft  $A$  und  $B$ , unabhängig voneinander erhoben und die Stetigkeit der erhobenen Merkmale der Objekte, sowie für beide Stichprobe gleich viele Landmarken

$$H_0 : P = Q$$

Die mittlere Abweichung der Objekte bei der Eigenschaft  $A$  hat die gleiche Verteilung wie bei Eigenschaft  $B$

$$H_1 : P \neq Q$$

**1.Schritt: Berechnung der mittleren Form:** Man berechnet die mittlere Form der ersten Gruppe mit Eigenschaft  $A$ . Die erste Gruppe besteht aus  $r$  von  $n$  Objekten. Erhalte Landmarken der mittleren Form  $l_i^m$ .

**2.Schritt: Berechnung des mittleren Abstandes:** Man berechnet die mittlere Abweichung  $m_i$  der  $i$ -ten Landmarke von der  $i$ -ten Landmarke der mittleren Form bei  $k$  Landmarken jeweils für die erste und zweite Gruppe mit jeweils Eigenschaft  $A$  oder  $B$ .

$$m_i^A = \frac{1}{r} \sum_{j=1}^r (l_i^j - l_i^m)^2$$

und

$$m_i^B = \frac{1}{n-r} \sum_{j=r+1}^n (l_i^j - l_i^m)^2$$

**3.Schritt: Vergabe der Ränge:** Man bestimmt die Differenz zwischen

den mittleren Abweichung aus  $A$  und  $B$  und gebe dem Betrag der Differenz den Rang. Große Differenz bedeutet hoher Rang.

**4.Schritt: Vergleich der Ränge für die jeweilige Gruppe:** Ist die Differenz zwischen der mittleren Abweichung der Landmarke  $i$  aus  $A$  und  $B$  positiv, so werden die Ränge aller darunter liegenden Differenzen addiert. Erhalte  $u_0^+$  analog zum “Wilcoxon-Test”.

**5.Schritt: Bewertung des Unterschieds zwischen den Gruppen:** Man vergleicht  $u_0^+$  mit allen anderen Möglichkeiten  $\binom{n}{r}$ ,  $r$  aus  $n$  Objekten zu ziehen. Der Rang von  $u_0^+$  wird in dieser Menge gebildet und gesetzt dem Fall, dass jeder Rang gleich wahrscheinlich ist ergibt sich ein  $p = \frac{\text{rang}(u_0^+)}{\binom{n}{r}}$  dafür, dass es keinen Unterschied bei der mittleren Abweichung für die Landmarken gibt.

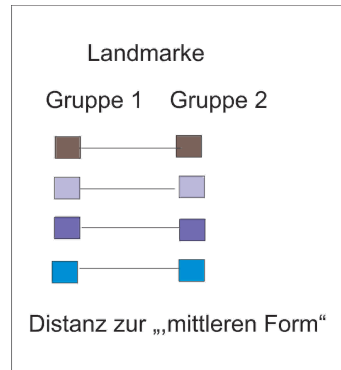


Abbildung 4.5: Wilcoxon-Test: Vergleich der Landmarken der Gruppe 1 mit Gruppe 2 über die Distanz zu der mittleren Form einer der beiden Gruppen

**Beispiel 24.** Geht man das oberhalb beschriebene Verfahren für die in der Pilotstudie [32] zur Verfügung stehende Kontingent an Objekten durch (Gruppe der 18 Wilms- Tumore zu den 3 Neuroblastomen) so erhält man einen  $u_0^+$ -Wert von 79. D.h. die mittlere Form wird aus der Gruppe der Wilms- Tumore berechnet. Der dem  $u_0^+$  entsprechende Wert kommt 5 mal vor und 250 Werte liegen unterhalb des Wertes bei insgesamt  $\binom{21}{3} = 1330$  Möglichkeiten. Entsprechend liegt der  $p$ - Wert im Intervall  $[\frac{251}{1330} = 0.1887, \frac{255}{1330} = 0.1917]$ . Für das Verwerfen von  $H_0$ , dass keinerlei Unterschied besteht, ist bei zweiseitigem Test ein  $\alpha > 2 \cdot 0.1917 = 0.3834$  zu wählen.

Andersherum für die mittlere Form der Neuroblastome ergibt sich ein  $u_0^+$ - Wert von 71. 1008 Werte liegen unterhalb und der Wert kommt 7 mal vor,  $p \in [\frac{1009}{1330} = 0.7586, \frac{1015}{1330} = 0.7632]$ . In diesem Fall wäre ein  $\alpha > 2 \cdot (1 - 0.7632) = 0.4736$  zu

---

wählen.

In der damals vorliegenden Erhebung erscheint die Vergabe des Rangs durch die Abweichung von der mittleren Form bei den Landmarken in der Gruppe von Patienten mit Diagnose Wilms oder Neuroblastom nicht auszureichen, um von einem signifikanten Unterschied bei den beiden Gruppen nämlich der Wilms-Tumore und der Neuroblastome bezüglich ihrer Abweichung zur mittleren Form sprechen zu können.

Tabelle 4.1: Abstand der Landmarken von den Landmarken der mittleren Form der Wilms- Tumore

Landmarke	Abstand			
Landmarke	$\bar{d}_{landmarke}^{wilms}$	$rang^{wilms}$	$\bar{d}_{landmarke}^{neuro}$	$rang^{neuro}$
Nr. 1	0.0063	11	0.0214	16
Nr. 2	0.0053	3	0.0271	21
Nr. 3	0.0064	13	0.0271	20
Nr. 4	0.0063	9	0.0187	14
Nr. 5	0.0068	16	0.0146	8
Nr. 6	0.0061	6	0.0135	7
Nr. 7	0.0055	4	0.0051	1
Nr. 8	0.0064	14	0.0103	2
Nr. 9	0.0061	7	0.0181	12
Nr. 10	0.0065	15	0.0219	17
Nr. 11	0.0071	18	0.0170	10
Nr. 12	0.0082	23	0.0125	5
Nr. 13	0.0090	24	0.0125	4
Nr. 14	0.0078	20	0.0191	15
Nr. 15	0.0063	12	0.0184	13
Nr. 16	0.0057	5	0.0131	6
Nr. 17	0.0072	19	0.0124	3
Nr. 18	0.0063	10	0.0178	11
Nr. 19	0.0051	2	0.0238	18
Nr. 20	0.0080	22	0.0253	19
Nr. 21	0.0044	1	0.0355	22
Nr. 22	0.0062	8	0.0154	9
Nr. 23	0.0068	17	0.0372	23
Nr. 24	0.0078	21	0.0738	24

Tabelle 4.2: Abstand der Landmarken von den Landmarken der mittleren Form der Wilms- Tumore

Landmarke	Abstand			
Landmarke	$\bar{d}_{landmarke}^{wilms}$	$rang^{wilms}$	$\bar{d}_{landmarke}^{neuro}$	$rang^{neuro}$
Nr. 1	0.0063	11	0.0214	16
Nr. 2	0.0053	3	0.0271	21
Nr. 3	0.0064	13	0.0271	20
Nr. 4	0.0063	9	0.0187	14
Nr. 5	0.0068	16	0.0146	8
Nr. 6	0.0061	6	0.0135	7
Nr. 7	0.0055	4	0.0051	1
Nr. 8	0.0064	14	0.0103	2
Nr. 9	0.0061	7	0.0181	12
Nr. 10	0.0065	15	0.0219	17
Nr. 11	0.0071	18	0.0170	10
Nr. 12	0.0082	23	0.0125	5
Nr. 13	0.0090	24	0.0125	4
Nr. 14	0.0078	20	0.0191	15
Nr. 15	0.0063	12	0.0184	13
Nr. 16	0.0057	5	0.0131	6
Nr. 17	0.0072	19	0.0124	3
Nr. 18	0.0063	10	0.0178	11
Nr. 19	0.0051	2	0.0238	18
Nr. 20	0.0080	22	0.0253	19
Nr. 21	0.0044	1	0.0355	22
Nr. 22	0.0062	8	0.0154	9
Nr. 23	0.0068	17	0.0372	23
Nr. 24	0.0078	21	0.0738	24

---

Anhand der Tabellen kann festgestellt werden, welche der Landmarken besonders von den Landmarken der mittleren Form abweichen und möglicherweise sich zur Differenzierung der Objektgruppen eignen.

---



## Kapitel 5

# Funktionale Formanalyse

### 5.1 Mittlere zu erwartende Funktion

Die Bestimmung einer mittleren zu erwartenden Funktion kann in unterschiedlichen Disziplinen von Interesse sein, um den generellen Trend aus mehreren den Trend beschreibenden Verläufen zu erkennen.

In der Finanzmathematik ist die Marktentwicklung von Rohstoffen, Aktien oder eines Portfolios zu beschreiben. Klassisches Beispiel sind die Marktindizes. Dabei ergibt sich die Einschätzung der Marktentwicklung aus der Bedeutung bzw. Gewichtung der einzelnen Verläufe für den Gesamtverlauf. Ein kleineres Unternehmen trägt in der Regel weniger bei als ein grösseres Unternehmen, um den Markttrend zu erkennen. Der Ansatz mehrerer je nach Bedeutung, Anteil am Marktvolumen etc. zu gewichtender Funktionen, um den generellen Trend zu erkennen, ist als *Basket* bekannt.

Die Bedeutung der einzelnen Verläufe wird in Form der Gewichte  $w_i$  für  $i = 1 \dots n$  dargestellt, mit denen die einzelnen  $n$  Verläufe multipliziert werden. Die stochastischen Funktionen ergeben sich als Erwartung aus den zu untersuchenden Verläufen  $S_i$ , somit  $f_i = E[S_i]$ . Die gewichtete mittlere Funktion ergibt sich zu:  $m_w = \sum_{i=1}^n w_i \underbrace{S_i(X)}_{\text{stochastische Funktionen}}$ .

Der Erwartungswert:  $E[m] = \sum_{i=1}^n w_i f_i$ . Bei gleicher Bedeutung, Marktvolumen etc. wird in der Regel als Gewicht  $w_i = 1$  gewählt.

**Beispiel 25.** Wir untersuchen drei historische Rohstoffurse vom 19. April 2010 bis zum 30. April 2010, Benzin(1), Palladium(2) und Bauholz(3). Zunächst wird die Lagrange-Interpolation verwendet. Es ergeben sich folgende Funktionen:

$$f_1(x) = 7.7436 - 14.8926x + 16.0330x^2 - 9.1528x^3 + 3.1014x^4 - 0.6546x^5 + 0.0869x^6 - 0.0071x^7 + 0.0003x^8 - 6.2354 \cdot 10^{-6}x^9$$

---


$$f_2(x) = -466 + 2819.7532x - 3203.6382x^2 + 1960.1778x^3 - 714.3407x^4 + 161.9747x^5 - 23.02528x^6 + 1.9924x^7 - 0.0958x^8 + 0.0020x^9$$

$$f_3(x) = 32 + 858.0226x - 961.7243x^2 + 565.3326x^3 - 195.1682x^4 + 41.6865x^5 - 5.5865x^6 + 0.4579x^7 - 0.0210x^8 + 0.00041x^9$$

Die mittlere zu erwartende Funktion bei den drei historischen Rohstoffkursen lautet:

$$m(x) = -142.0855 + 1220.9611x - 1383.1098x^2 + 838.7859x^3 - 302.1358x^4 + 67.6689x^5 - 9.5083x^6 + 0.8144x^7 - 0.0388x^8 + 0.0008x^9.$$

Sie entspricht dem Mittelwert, da keine Rotationen bei der Bestimmung der mittleren zu erwartenden Funktion zu beachten sind.

**Beispiel 26.** Bei einem Portfolio aus drei Aktienkursen von Linde(1), Luft-hansa(2) und Deutsche Bank(3) ist die Gewichtung der jeweiligen Funktion zu verändern. Das Verhältnis beträgt:  $m_1 : m_2 : m_3$ . Die Funktion  $f_1$  wird dann mit  $\frac{m_1}{m_1+m_2+m_3}$ ,  $f_2$  mit  $\frac{m_2}{m_1+m_2+m_3}$  und  $f_3$  mit  $\frac{m_3}{m_1+m_2+m_3}$  gewichtet. Im rechnerischen Beispiel ist das Verhältnis 2 : 1 : 3. Die mittlere Funktion lautet:

$$m(x) = -107.2744 + 464.0132x - 523.6490x^2 + 318.1051x^3 - 115.6415x^4 + 26.2599x^5 - 3.7472x^6 + 0.3259x^7 - 0.0158x^8 + 0.0003x^9.$$

Der Algorithmus von Ziezold(1994)[90] berechnet eine mittlere Form im Zweidimensionalen. Wie bereits erwähnt ist der arithmetische Mittelwert bei mehr als zwei Objekten nicht unbedingt gleich dem Erwartungswert nach Fréchet. Im bisherigen Ansatz einer Berechnung eines gewichteten Mittelwertes wird diese Erkenntnis ignoriert.

Die erstmalige Übertragung des Algorithmus von Ziezold(1994)[90] ermöglicht eine Beachtung der im Zweidimensionalen vorliegenden Rotationen.

Der Algorithmus von Ziezold kann auf Funktionen im  $\mathbb{R}^2$  und  $\mathbb{R}^3$  angewandt werden. Falls nur einzelne Punkte bekannt sind und nicht die gesamte Funktion, benutzt man Interpolationen, um die Funktionen zu schätzen.

Das im Algorithmus angewandte Skalarprodukt  $\langle \cdot, \cdot \rangle$  ist das Skalarprodukt  $\langle f, g \rangle = \int_{-\pi}^{\pi} f \bar{g} dx$ , wobei  $\bar{g}$  die komplex konjugierte Funktion von  $g$  ist.

Die mittlere Funktion ergibt sich dann nach unserem erstmalig auf Funktionen angewandten Algorithmus:

---

1. Für  $i = 1, \dots, n$  und die Funktionen  $f_i$ :

$$\tilde{m} \mapsto w_i(\tilde{m}) = \begin{cases} \frac{\langle \tilde{m}, f_i \rangle}{|\langle \tilde{m}, f_i \rangle|} & \text{wenn } \langle \tilde{m}, f_i \rangle \neq 0 \\ 1 & \text{wenn } \langle \tilde{m}, f_i \rangle = 0 \end{cases} \quad (5.1)$$

2.

$$\tilde{m} \mapsto T(\tilde{m}) = \frac{1}{n} \sum_{i=1}^n w_i(\tilde{m}) f_i \quad (5.2)$$

3. Dann berechnen wir rekursiv

$$\tilde{m}_r = T(\tilde{m}_{r-1}), r = 1, 2, \dots, \quad (5.3)$$

4. bis das folgende Abbruchkriterium erfüllt ist:

$$\tilde{m} = T(\tilde{m}) \quad (5.4)$$

**Beispiel 27.** Seien die drei Funktionen

$$\begin{aligned} f_1(x) &= -e^{x^2} + 2x - 2x^2 \\ f_2(x) &= 5x^3 - x^4 \\ f_3(x) &= -1 + x^5 + x^7. \end{aligned}$$

Die mittlere zu erwartende Funktion ist dann

$$m(x) = -0.3333 - 0.3333e^{x^2} + 0.6667x - 0.6667x^2 + 1.6667x^3 - 0.3333x^4 + 0.3333x^5 + 0.3333x^7.$$

**Beispiel 28.** In der von mir durchgeführten "Rückfalluntersuchung im rheinland-pfälzischen Jugendstrafvollzug von 1996 bis 2000" wurden 400 jugendliche Straftäter erfasst[47]. Es wird zur Bestimmung der "Rückfälligkeit" in Form eines Freiheitsentzugs (57.5% ) der Heimwechsel  $x$  (0:"nein", 1:"ja"), das Alter  $y$  (14 bis 25) und das begangene Gewaltdelikt  $z$  (0:"nein", 1:"ja") verwendet.

Es wird die Lineare Regression und die Logistische Regression angewendet:

$$f_1(x) = 0.569 + 0.230x - 0.001y - 0.037z.$$

$$f_2(x) = \frac{1}{1 + e^{-(0.292 + 1.030x - 0.006y - 0.158z)}}$$

Die mittlere Funktion aus beiden Funktionen entspricht bei gleicher Gewichtung dem Mittelwert. Bei mehr als zwei Funktionen ist das Skalarprodukt entsprechend zu erweitern:  $\int_{z_{min}}^{z_{max}} \int_{y_{min}}^{y_{max}} \int_{x_{min}}^{x_{max}} fg \, dx dy dz$ . Die Gewichtung der Modelle kann auch nochmalig geschätzt werden.

**Beispiel 29.** Bei drei dreidimensionalen Körpern  $(x, y, z \in \mathbb{R})$  können zwei Dimensionen im Komplexen zusammengefasst werden.

$$f_1(x) = (1 - 2ix + x^2)(1 + 2ix + x^2).$$

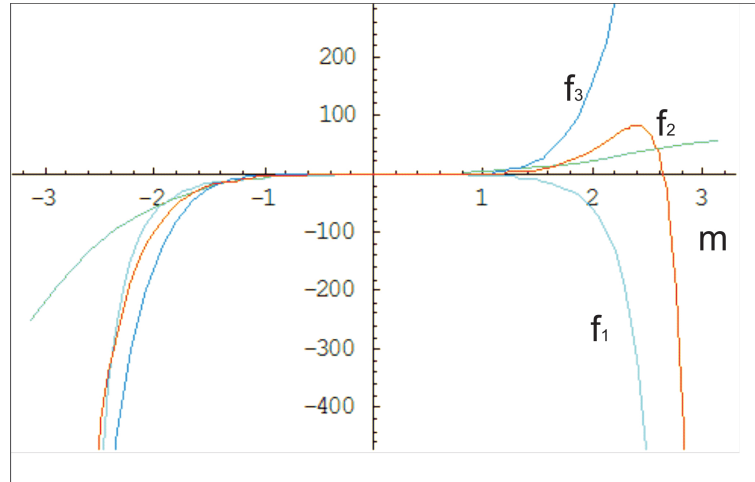


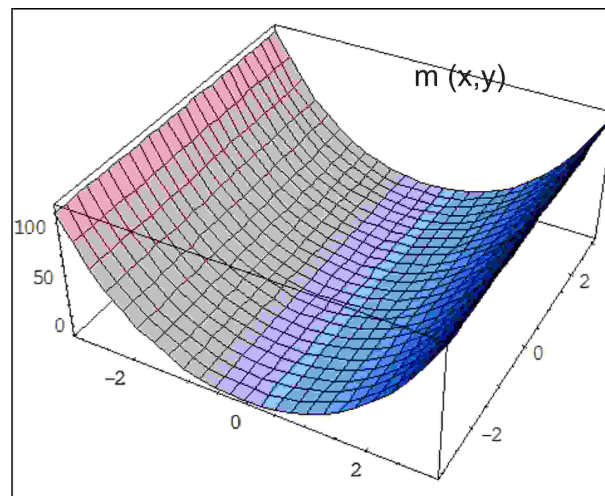
Abbildung 5.1: Mittlere Funktion von drei reellen Funktionen

$$f_2(x) = (4ix + x^4)(-4ix + x^4).$$

$$f_3(x) = (1 + 2ix + x)(1 - 2ix + x).$$

Die mittlere Form der drei Funktionen ist

$$m(x) = 0.6667 + 2.6946 \cdot 10^{-28}i + 0.6667x + (9 + 1.6186 \cdot 10^{-27}i)x^2 + (0.3333 + 2.6946 \cdot 10^{-28}i)x^4 + (0.3333 + 1.2262 \cdot 10^{-31}i)x^8.$$



Mittlere Form von drei Funktionen

## 5.2 Dynamische Formanalyse

Bisher völlig unbeachtet auf dem Gebiet der Formanalyse sind die Formveränderungen in der Zeit. Diese Veränderungen der Form innerhalb der Zeit werden als neues Gebiet der dynamischen Formanalyse erstmalig in dieser Arbeit angegangen. Eine stetige Betrachtung der Veränderung der Form in der Zeit ist angesichts der für die Anwendung häufig hervorgehobenen Bedeutung bestimmter Zeitpunkte gegenüber allen anderen Zeitpunkten und einer zumeist fehlenden stetigen Messung als nicht sinnvoll anzusehen. Darüber hinaus ergibt sich die Vergleichbarkeit der Formveränderungen in der Zeit häufig erst durch die anwendungsspezifisch vorgegebenen Zeitpunkte.

Zu vergleichbaren Zeitpunkten  $t = 1 \dots T$  werden die Landmarken erhoben,  $t > 1$ . Die Vergleichbarkeit der Zeitpunkte kann sowohl die Messung zu einem gleichen Messzeitpunkt bedeuten, als auch zu einem Zeitpunkt, in dem die zu untersuchende und damit zu vergleichende Veränderung gemessen werden kann:  $\forall i = 1 \dots n, \forall t = 1 \dots T, o_{it} = \{l_1^t \dots l_k^t\}, l_j^t \in \mathbb{R}^m$ .

Für jedes Objekt wird von jedem diskreten Zeitpunkt  $t$  zum nächsten Zeitpunkt  $t + 1$  eine Funktion  $g$  ermittelt, so dass  $g(o_i^t) = o_i^{t+1}$  gilt. Bei  $k$  Landmarken kann dafür ein lineares Gleichungssystem mit  $k$  Gleichungen angesetzt werden. Unter der Annahme, dass für die Veränderung des  $i$ -ten Objekts bei der  $j$ -ten Landmarke nur die Veränderung der  $j$ -ten Landmarke zu dem darauf folgenden Zeitpunkt zu betrachten ist ergibt sich für jede Landmarke von dem Zeitpunkt  $t$  zum Zeitpunkt  $t + 1$  die Gleichung mit den Koeffizienten  $w_{j,r}^{it}, \forall r = 1, \dots, m$ :

$$l_j^{i^{t+1}} = W_j^{it} l_j^{it}, W_j^{it} \in \mathbb{R}^{m \times m} \quad (5.5)$$

Für die Diagonale der Matrix ergibt sich, wenn  $l_{j,r}^{it} \neq 0$ :

$$w_{j,r}^{it} = \frac{l_{j,r}^{i^{t+1}}}{l_{j,r}^{it}}, \forall r = 1, \dots, m \quad (5.6)$$

Wenn  $l_{j,r}^{it} = 0$  bzw. bei allen Nicht-Diagonaleinträgen, ist  $w_{j,r}^{it} = 0$  bei diesem Ansatz.

Anstatt die Distanz zur "mittleren Form" wird die Distanz der Veränderung des  $i$ -ten Objektes zu der Veränderung der "mittleren Form" verwendet. Betrachtet wird die jeweils geeignet zu definierende Distanz zwischen der Veränderung einer Form zwischen zwei Zeitpunkten  $o^{it}$  und  $o^{i^{t+1}}$  und der Veränderung der zu den jeweiligen Zeitpunkten zu berechnenden "mittleren Form"  $\tilde{m}^t$  und  $\tilde{m}^{t+1}$  aller gemessenen Objekte zu den zu untersuchenden Zeitpunkten. Dafür kann die Differenz der Koeffizienten  $w_{j,r}^{it}$  und der der mittleren Form  $w_{j,r}^{\tilde{m}}$  verwendet werden.

Um zu einem reellen Wert  $d_{o^i, \tilde{m}}$  zu gelangen, der den Unterschied zwischen der Veränderungen des  $i$ -ten Objekts und der der mittleren Form  $\tilde{m}$  beschreibt, kann eine geeignete Matrixnorm für jede der für die Landmarken bestimmte

---

$W_j^{i^t}$  gewählt und entsprechend aufsummiert werden. Dieser reelle Wert  $d_{o^i, \bar{m}}$  ist dann u.a. analog dem bereits beschriebenen Test von Ziezold(1994) [90] zu prüfen. Ebenfalls können die von uns bereits vorgenommenen Erweiterung des Tests für die Varianz alleine und unter Betrachtung von Varianz und mittlerer Form angewandt werden. Sollen mehr als nur zwei Zeitpunkte betrachtet werden sind die Unterschiede für jeweils zwei aufeinander folgende Zeitpunkte aufzusummieren. So ergeben sich für drei Zeitpunkte zwei zu betrachtende Veränderungen usw. Als Anwendung für die dynamische Formanalyse für zwei aufeinander folgende Zeitpunkte in der Medizin werden im Weiteren die Herzdaten verwendet.

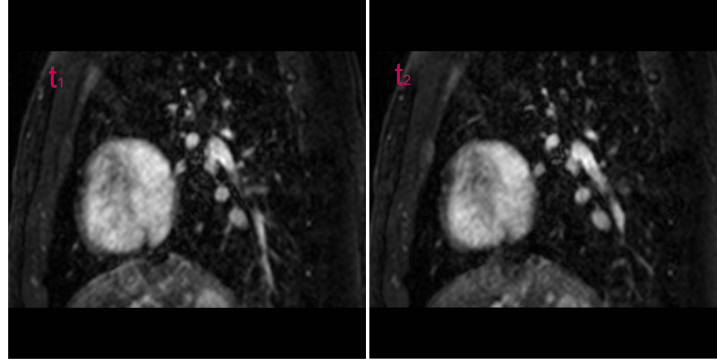


Abbildung 5.2: Herzsequenz: Herz zum Zeitpunkt  $t_1$  und zum Zeitpunkt  $t_2$

Um den zeitlichen Verlauf der Patienten vergleichen zu können, werden Zeitpunkte gewählt, die aufgrund ihrer medizinischen Bedeutung vergleichbar sind. Bei den Herzdaten sind die Zeitpunkte die Diastole und Systole. Im folgenden eine dreidimensionale Darstellung des Herzens zum Zeitpunkt der Diastole in der Abbildung 5.3.

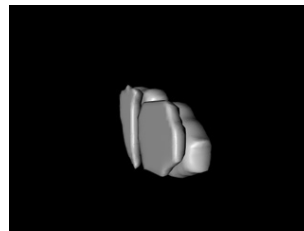


Abbildung 5.3: Herzsequenz: Herz zum Zeitpunkt der Diastole

---

**Beispiel 30.** Im Rahmen der Kooperation mit dem Universitätsklinikum Heidelberg ergaben sich bisher nur 5 kranke und 5 gesunde Herzen. Dabei ist zu beachten, dass die gesunden Herzen in das Patientenkollektiv kamen, weil ein Verdacht auf eine Herzerkrankung besteht und es sich damit um Herzen handelt, bei denen zumindest der Verdacht auf eine Erkrankung besteht. Ein Ergebnis der dynamischen Formanalyse ist erst bei einer entsprechenden Erhöhung der Fallzahl und einer Gewährleistung des Vergleichs zwischen gesunden und kranken Herzen zu erwarten. Für die Nierentumore ist eine solche Untersuchung nicht angedacht, da zum einen nur wenige Patienten zu unterschiedlichen Zeitpunkten eine MRT-Aufnahmen hatten und zum anderen hierfür ein unterschiedliches Wachstumsverhalten an sich bei den Tumorarten angenommen werden müsste und entsprechend für eine solche Annahme erhoben werden.

---



## Kapitel 6

# Neuronale Netze

Der erste Gedanke der neuronalen Netze ist es, die Verarbeitung innerhalb der Nervenzellen mathematisch zu modellieren. Die Eingabe in Form von chemischen, optischen und mechanischen Reizen ist so zu verarbeiten, dass eine entsprechende Reaktion der Nervenzelle gegenüber anderen Zellen erfolgen kann[55]. Die Ausgabe des Netzes kann zum einen in Form einer geeigneten Reaktion gegenüber den anderen Zellen vorgegeben sein und zum anderen in Form einer Informationsverarbeitung und Kategorienbildung sich aus den eingegebenen Daten von selbst ergeben. Der erste Ansatz ist als überwachtes Netz bekannt, der zweite Ansatz als unüberwachtes Netz in Form der Kohonenkarte [74].

Der Begriff "neuronal" suggeriert ein selbstlernendes Verfahren, jedoch ist bei beiden Ansätzen sowohl bei der Bestimmung einer Zuordnung zwischen Ein- und gewünschter Ausgabe, als auch bei der Clusterbildung ähnlichen Kohonenkarte zu erkennen, dass es sich um ein Verfahren handelt, das durch die mathematischen Vorgaben (Wechselwirkungen zwischen allen Eingabevariablen) klar eingeschränkt ist. Das Ziel des Netzes ist es bei beiden Ansätzen ohne statistische Verteilungsannahmen etc., eine zu betrachtende Distanzfunktion durch einen das Lernen beschreibenden Algorithmus zu minimieren. Neuronale Netze sind somit ursprünglich ein numerisches Verfahren.

### 6.1 Elemente Neuronaler Netze

Zunächst sind sie entwickelt worden, um die kognitiven Prozesse zu verstehen. Heutzutage gibt es eine Vielzahl an Anwendungen Neuronaler Netze als numerisches Verfahren. .

---

## Das Neuron:

McCulloch und Pitts modellierten die biologischen Prozesse einer Nervenzelle zum allerersten Mal in einer mathematischen Form [55]. Nervenzellen haben in diesem Ansatz die Daten aus ihrer Umgebung aufzunehmen und zu verarbeiten, um bestimmte Zielgrößen bestimmen zu können.

Die Daten setzten sich bei den ersten Anwendungen aus  $n$  Variablen mit  $x_1, \dots, x_n$  auf binärem Messniveau zusammen und dienten zur Bestimmung einer binären Zielgröße.

Für die Verarbeitung der Daten wird die  $i$ -te Variable  $x_i$  gewichtet mit  $w_i$ . Wenn die Gewichtungen im Betrag betrachtet werden, kann aus dem Betrag  $|w_i|$  die Bedeutung der standardisierten Eingangsvariablen für die Bestimmung der Zielgröße bei einem einschichtigen Neuronalen Netz bestimmt werden. Aus dem Vorzeichen ergibt sich die Richtung des Einflusses. Die Gewichtung der Eingangsvariable ist mit der der Diskriminanzanalyse vergleichbar.

Der kritische Wert für das Neuron ist die gewichtete Summe der Eingangsvariablen:

$$q := \sum_{i=1}^n w_i \cdot x_i = w_1 \cdot x_1 + \dots + w_n \cdot x_n \quad . \quad (6.1)$$

Für eine Zielgröße  $y$  mit binärem Messniveau wird ein Schwellenwert  $S$  benötigt. Das Überschreiten des Schwellenwertes führt zu  $y = 1$  und das Unterschreiten zu  $y = 0$ . Daher wird eine Aktivierungsfunktion  $F$  auf die gewichtete Summe der Eingangsvariablen angewandt:

$$F(q) = \begin{cases} 1, & \text{wenn } x > S \\ 0, & \text{wenn } x \leq S \end{cases} \quad (6.2)$$

Im Vergleich zur Diskriminanzanalyse ist der Schwellenwert  $S$  vorzugeben, abhängig von den Eigenschaften der Zielgröße. Der Schwellenwert wird nicht aus den Daten hergeleitet. Neuronale Netze enthalten zur Schätzung eines solchen Schwellenwertes keine Annahmen über die Daten und ihre Verteilung. Mit der Anwendung der Bewertungsfunktion erhalten wir  $y = F(q)$ :

$$\begin{aligned} y &= 1, & \text{if } \sum_{i=1}^n w_i \cdot x_i > S \\ y &= 0, & \text{if } \sum_{i=1}^n w_i \cdot x_i \leq S \end{aligned}$$

Als Aktivierungsfunktion wird zumeist die Sigmoidfunktion verwendet. Ihre Werte lassen sich nicht nur wie in den bisherigen Betrachtungen nur bezüglich des Schwellenwertes deuten, sondern bei einer entsprechenden Einschränkung des Wertebereichs  $[0, 1]$  und einem Schwellenwert von 0.5 als Wahrscheinlichkeiten für die Zuordnung zu einer der beiden Kategorien deuten. Während ein

---

einschichtiges neuronales Netz die logische Funktion "OR" darstellen kann, ist sie nicht in der Lage, die "XOR"-Funktion darzustellen [77]. Die im nächsten Abschnitt beschriebenen mehrschichtigen Neuronale Netze sind zur Abbildung aller logischen Zuordnungen in der Lage.

## 6.2 Mehrschichtiges neuronales Netz

Im Allgemeinen ist die Bestimmung einer Zielgröße im Laufe des Lernens des Netzes mit einem zu minimierenden Fehler verbunden. Ein gegebenes Maß für den Fehler  $E(\tilde{y}, y)$  ist der Abstand zwischen Zielgröße und dem vom Neuronalen Netz berechneten Wert. Das Lernen des Neuronalen Netzes korrespondiert mit der Minimierung des Fehlers für  $E(\tilde{y}, y)$ . Der folgende Algorithmus geht auf Rumelhart, Hinton und Williams [64] zurück. Der Gesamtfehler über alle Werte ist definiert als:

$$E_{total}(\tilde{y}, y) := \frac{1}{2} \sum_{k=1}^N (\tilde{y}_k - y_k)^2 \quad . \quad (6.3)$$

Der Gesamtfehler wird benutzt, um die Gewichte in jeder Schicht des Neuronalen Netzes zu rekallibrieren. Der Einfachheit halber wird ein 2-Schichten Netzwerk verwendet, welches ausreicht alle logischen Funktionen darzustellen. Der Wert  $\tilde{y}$  des Neuronalen Netzes wird berechnet in folgenden Schritten. Zunächst wird der Parameter für die erste Schicht berechnet anhand der  $n$  gewichteten Eingangsvariablen  $\sum_{i=1}^n w_i \cdot x_i$ . Wir betrachten eine versteckte Schicht mit  $m$  Neuronen. Für  $j = 1, \dots, m$ , sei  $g_j$  die Aktivierungsfunktion im  $j$ -ten Neuron der versteckten Schicht, mit einem Wert von  $h_j$ , gegeben als

$$h_j = g_j\left(\sum_{i=1}^n w_i \cdot x_i\right) \quad . \quad (6.4)$$

Gewöhnlich wird für alle Neuronen einer gegebenen Schicht eine Aktivierungsfunktion  $g = g_1, \dots, g_m$ , wie bspw. die Sigmoidfunktion, verwendet. Als nächstes wird die Ausgabe der vorherigen versteckten Schicht die Eingabe der nächsten Schicht. Sei  $f$  die Aktivierungsfunktion der vorletzten Ausgabeschicht, dann ist der entsprechende Wert der vorletzten Schicht definiert als:

$$q = f\left(\sum_{j=1}^m u_j \cdot h_j\right) \quad . \quad (6.5)$$

Abschließend wird der Ausgabewert des Netzes  $q$  bewertet anhand der Funktion  $F$  :

$$\tilde{y} = F(q) \quad (6.6)$$

als endgültige Ausgabe des Netzes anhand der Gewichtung der Eingabevariablen und der versteckten Neuronen erhalten wir  $\tilde{y}$ . Das Neuronale Netz

---

verändert nun in jeder Iteration die Gewichte. Der Lernmechanismus der Gewichte wird bestimmt durch den Abstand der Ausgabe des Neuronalen Netzes zur Zielgröße.

$$E = \frac{1}{2} \sum_{i=1}^n (y^i - \tilde{y}^i)^2 \quad .$$

Die Gewichte in beiden Schichten werden nach dem Gradientenverfahren verändert, i.e.

$$\nabla w_i = \frac{\partial E}{\partial w_i} \quad (6.7)$$

$$\nabla u_j = \frac{\partial E}{\partial u_j} \quad (6.8)$$

Die Lernrate  $\alpha$  sollte auf die Daten und das zu lösende Problem angepasst sein. Die Gewichte werden nach folgenden Regeln verändert:

$$w_i^{new} = w_i^{old} - \alpha \cdot \nabla w_i \quad (6.9)$$

$$u_j^{new} = u_j^{old} - \alpha \cdot \nabla u_j \quad (6.10)$$

Die notwendige Anzahl an Iterationen hängt von den Anforderungen der Daten, des Anwenders und der Disziplin ab. Das hier vorgestellte mehrschichtige Neuronale Netz kann dafür verwendet werden, die Wahrscheinlichkeit  $p$  für die Diagnose zu bestimmen.

Es ist in seiner einschichtigen Form in Folge der Verwendung der nichtlinearen Funktionen und bei entsprechender Einschränkung des Wertebereichs mit der Logistischen Regression vergleichbar. Lediglich die Bestimmung der Gewichte erfolgt auf numerischen Weg und nicht anhand von Verteilungsannahmen.

### 6.3 Neuronale Netze in Kombination mit stochastischen Prozessen

Neuronale Netze werden häufig als ein Verfahren angesehen, dessen Rechenprozesse nicht offen gelegt werden können. Gerade in mehrschichtigen Neuronalen Netzen werden die Rechenprozesse durch die Vernetzung der Neuronen zu komplex. Ein möglicher Ansatz zur Offenlegung ist es, über eine veränderte Eingabe die Ausgabe des Netzes zu kontrollieren und damit die Bedeutung der einzelnen Eingaben zu ermitteln [22].

In Neuronalen Netzen können im Weiteren bereits vorhandene Kenntnisse über die Interaktionen der eingegebenen Variablen etc. integriert werden. Sind dem Anwender bestimmte Interaktionen oder Abhängigkeiten der erklärenden Merkmale bekannt, so ist es sinnvoll zur Beachtung dieser Interaktion diese bereits

---

durch entsprechende Vernetzung der Neuronen zu integrieren.

Ähnlich ist auch die Übertragung der stochastischen Prozesse und Ansätze in Neuronale Netze möglich. Die Neuronen bestimmen dann die für den stochastischen Prozess notwendigen Parameter. So kann ein Neuron den Mittelwert anhand der unterschiedlich gewichteten Eingangswerte bestimmen und ein anderes die Varianz [36].

**Beispiel 31.** Das numerische Verfahren der Neuronalen Netze kann durchaus kombiniert werden mit stochastischen Modellannahmen. Ein Neuronales Netz, kallibriert durch ein stochastisches Modell, wird angewandt zur Vorhersage des nächsten Tages. Im Folgenden verwenden wir einen festen bekannten Zeitraum (bspw. 10 Geschäftstage). Das Modell lernt die Gewichtungen anhand der historischen Daten. Mit den Gewichtungen werden die Parameter des stochastischen Modells optimiert, und der nächste Tag für den Gaspreis berechnet. Das Gedächtnis des Neuronalen Netzes ist begrenzt durch die verwendeten historischen Daten. Die Methode von Rainer und Giebel (2009) [36] wurde getestet mit dem Gaspreis von Nordpool für den Zeitraum vom 1.1.2009 bis 31.12.2009. Bei jeder Vorhersage wurden  $n = 10$  vorhergehende Werte für die Vorhersage des nächst folgenden Wertes verwendet. In der Grafik 6.1 sind die mit der Wirklichkeit zu vergleichenden vorhergesagten Werte dargestellt.

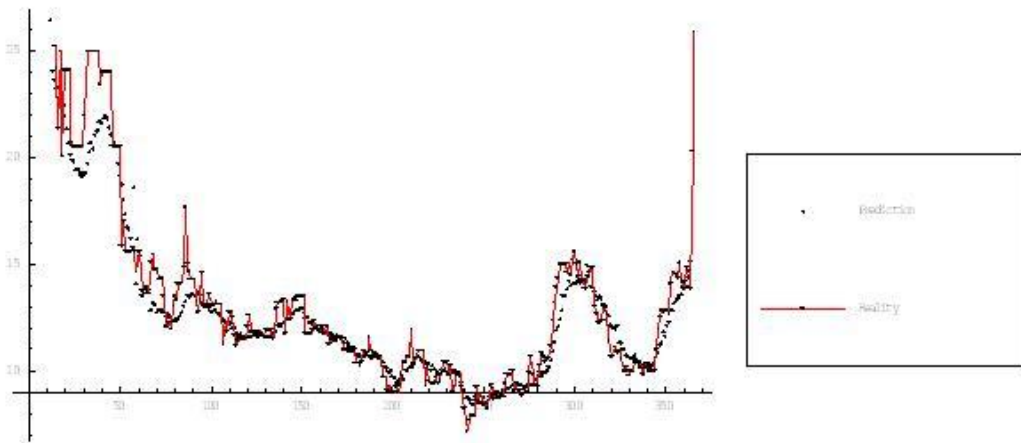


Abbildung 6.1: Gaspreis von Nordpool 2009

Direkt angewandt wurde dieses Vorgehen ebenfalls auf Klima- und Energiedaten [43] [42] [40].

---

## 6.4 Neuronale Netze in Kombination mit der Formanalyse

In dem oberen Ansatz von Rainer und Giebe(2010) [36] wurde die Kombination der stochastischen Prozesse und der Neuronalen Netzen in der Finanzmathematik vorgestellt.

Der von uns in der Finanzmathematik vorgestellte Ansatz einer Minimierung durch das Netz dient auch in der Formanalyse zur Erweiterung der bisher lediglich betrachteten euklidischen Transformationen von Objekten (Reflektion, Rotation, Translation und Skalierung). Bisher unbeachtet bleiben dabei die nicht-euklidischen Transformationen. Weitergehend stellt sich die Frage der Klassifikation durch Neuronale Netze, da die Interaktionen zwischen den Landmarken nicht auszuschließen sind und damit die Annahme der Unabhängigkeit bei der Verwendung der Logistischen Regression oder der Diskriminanzanalyse verletzt ist.

Es gibt vier verschiedene Ansätze, die Neuronalen Netze in die Formanalyse einzubeziehen:

- Klassifikation relevanter Merkmale gemäß der bisherigen Verwendung neuronaler Netze
- Klassifikation anhand der Distanz zur mittleren Form und im Falle einer Begründung der Varianz der erklärenden Merkmale
- Bestimmung der mittleren Form  $\mu$  mit der geringsten Varianz innerhalb der Objektgruppe
- Bestimmung einer mittleren Form  $\mu^*$  einer Objektgruppe mit der besten Differenzierbarkeit zu einer anderen Objektgruppe

Die bisherige Verwendung überwachter Neuronaler Netze im ersten Fall lief darauf hinaus, anhand erklärender Merkmale eine Klassifikation vorzunehmen. Die Beschreibung der Ausgabe des Neuronalen Netzes als Wahrscheinlichkeit und die Deutung der Ergebnisse und der Relevanz der Eingangsvariablen anhand unterschiedlicher Ausgangslagen geht auf die von mir vorgenommene Anwendung auf Daten zur rheinland-pfälzischen Rückfalluntersuchung jugendlicher Straftäter von 1996 bis 2000 [22] zurück. Die Deutung der Ergebnisse erfolgt über die kontrollierte Veränderung einiger Merkmale oder eines Merkmals auf die Ausgabe des Neuronalen Netzes. Der zweite Fall entspricht dem Vorgehen von Rainer und Giebel [36] und kombiniert die Stochastik mit der Numerik. Die Parameter der Verteilung werden numerisch durch das neuronale Netz bestimmt und anhand der Verteilungsparameter wird die entsprechende Klassifikation vorgenommen. Im dritten Fall werden die erhobenen Landmarken in den drei betrachteten Dimensionen unterschiedlich gewichtet, um die Varianz innerhalb der Gruppe zu minimieren. Im vierten Fall wird zusätzlich zur Minimierung der Varianz der  $u$ -Wert berechnet. Es wird die Form gewählt, die eine optimale Differenzierung

---

zu der entsprechenden anderen Objektgruppe zulässt.

Die Minimierung der Varianz führt dazu, dass wenn die Varianz immer mehr die Null erreicht, dass der stochastische Prozess abbricht. Daher ist bei alleiniger Minimierung der Varianz eine Verringerung der Differenzierung hinsichtlich anderer Objektgruppen zu erwarten. Mittels der Formel von Tschebyscheff ergibt sich  $P(|X - m| \geq \epsilon) = 0, \forall \epsilon > 0$ . Für  $P(|x - m| > 0) = P(x \neq m) = 0$ , für das Gegenereignis  $P(x = m) = 1$ . Demzufolge bedeutet die Varianz gleich Null ein Abbrechen des stochastischen Prozesses, da nur noch ein Ereignis nämlich  $A = \{x = m\}$  eintreten kann. Bei einer diskreten Verteilung bedeutet das, dass alle beobachteten Objekte dem Objekt  $m$  mit 100%iger Wahrscheinlichkeit entsprechen.

#### 6.4.1 Neuronale Netze zur Differenzierung der Gruppen

Völlig neu ist unser Ansatz [36] zur Kombination stochastischer Modelle und Neuronaler Netze. Dieser Ansatz kann in die Formanalyse nahtlos übertragen werden. So dient im folgenden das Neuronale Netz zur Bewertung des Abstands zur mittleren Form, die einen Erwartungswert innerhalb der Gruppe von Objekten darstellt.

Der Abstand zur mehrdimensionalen Varianz einer Gruppe von Objekten ist gerade im Falle einer theoretischen Begründbarkeit bereits im Neuronalen Netz zu beachten. In der im Abschnitt 8.1 beschriebenen Anwendung auf Tumore ist davon auszugehen, dass Tumore, die sich nur in einem klar körperlich beschränkten Raum entwickeln können, eine geringere Varianz aufweisen. Die Annahme der unterschiedlichen Varianz kann dann bereits in das Modell integriert werden. Gerade bei dreidimensionaler Betrachtung ist die Varianz in allen drei Dimensionen sehr unterschiedlich zu bewerten. Während hinsichtlich parallel der Wirbelsäule ein erheblicher Spielraum besteht, ist der Spielraum senkrecht als geringer einzuschätzen durch die Begrenzung der Wirbelsäule und des Körpers. Für die gewählte Tumorart gibt es keinerlei theoretische Vorkenntnis der Varianz.

Im Neuronalen Netz wird für die entsprechende Klasse, Diagnose oder Gruppe eine Wahrscheinlichkeit berechnet, die als Grundlage für eine Zuordnung dient. Durch den Ansatz einer zu deutenden Wahrscheinlichkeit für die Klasse werden anhand der vorliegenden Daten und/oder anhand des Verfahrens nicht klar erkennbare oder nur schwierig zuzuordnende Objekte als solche auch erkannt, da die Wahrscheinlichkeit nahe an der Schwelle ihrer jeweiligen Zuordnung zu den Kategorien liegen wird.

Diese Version kann für weitere Klassen oder Diagnosen fortgesetzt werden. So kann für jede Klasse der Abstand der jeweiligen Landmarke zur Landmarke der mittleren Form für eine Wahrscheinlichkeit für die entsprechende Klasse verwendet werden.

Im hier verwendeten Neuronalen Netz könnte jede Koordinate zur Klassifika-

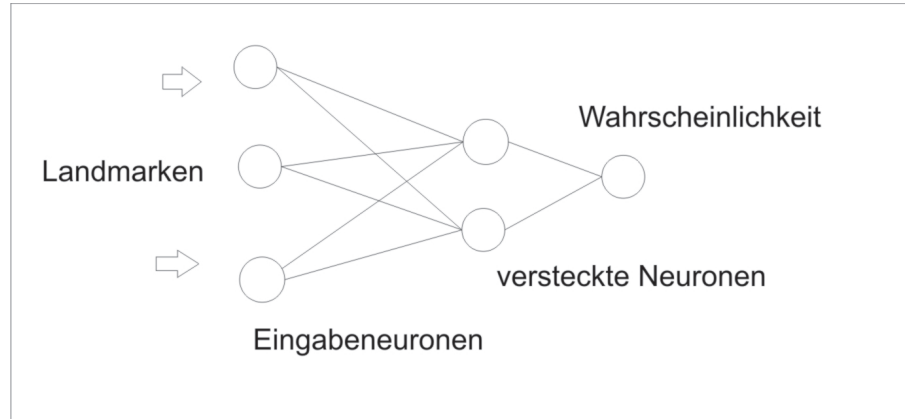


Abbildung 6.2: Neuronales Netz zur Ermittlung der Wahrscheinlichkeit für eine vorliegende Diagnose

tion verwendet werden. Dadurch würden sich 180 erklärende Variablen ergeben. Dies würde zwangsläufig bei einer geringeren Datenmenge als 180 Patienten zu einem Overfit des Neuronalen Netzes führen. Ein Overfit bedeutet, dass das Ergebnis für weitere Daten nicht verwendbar erscheint und die entsprechende Zuordnung zur Klasse auch durch ein zu lösendes Gleichungssystem bestimmt werden kann. Daher wird jede der 60 Landmarken als Summe ihrer Koordinaten zur Klassifikation verwendet.

Ein anderer Ansatz ist es, den Mittelwert der Distanzen und die Varianz durch das Netz zu bestimmen und damit analog zu dem bereits beschriebenen Vorgängen zu arbeiten. Die Distanzen zur "mittleren Form" werden in dem einen versteckten Neuron zur Bestimmung eines geeigneten Mittelwertes für die Distanz zur "mittleren Form" verwendet und im jeweils anderen versteckten Neuron zur Bestimmung der Varianz. Vorteil eines solchen Vorgehens ist, dass nicht jede Landmarke gleichermaßen gewichtet wird, sondern entsprechend ihrer Eignung zur Klassifikation der Tumore. Vor einer solchen Integration dieser Annahme wäre zunächst die unterschiedliche Varianz der Tumorarten weiter zu untersuchen und die Eignung eines solchen Mittelwertes. Die vorliegenden Daten berechtigen zwar zu Aussagen über die Nephroblastome, aber nur sehr bedingt über die Nicht-Nephroblastome.

Um die Neuronalen Netzen auch in ihrer üblichen Anwendung als Klassifikationsverfahren zu verwenden, wird zusätzlich zum Abstand zur mittleren Form der Nephroblastome das Alter des Patienten, die Lage des Tumors (Links, Rechts) gewählt.



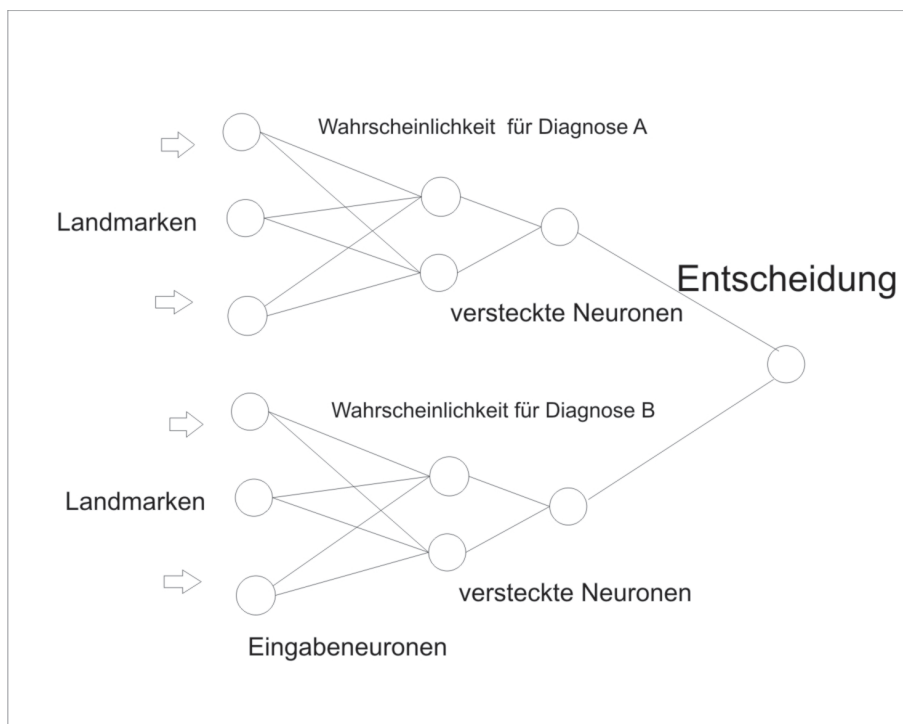


Abbildung 6.3: Neuronales Netz zur Ermittlung der Wahrscheinlichkeit für mehrerer vorliegende Diagnosen

---

# Kapitel 7

## Klassifikation

### 7.1 Einführung

Bei der Klassifikation der Objekte ist eine Zuordnung  $f$  zu finden, die entweder direkt das Objekt der jeweiligen Gruppe zuordnet oder mit einem deutbaren Wert wie die Wahrscheinlichkeit den Anwender in die Lage versetzt, diese Zuordnung ab einem bestimmten Wert vorzunehmen.

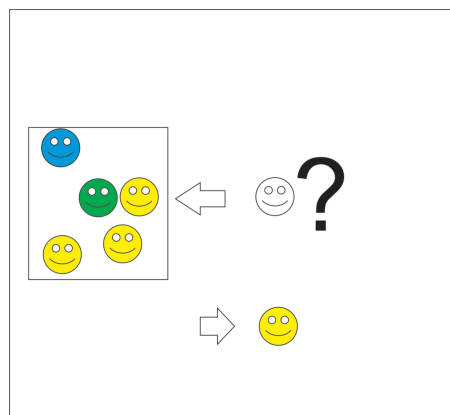


Abbildung 7.1: Grundgedanke der Klassifikation

Der Grundgedanke der Klassifikation ist es, eine noch unbekannte Person, Fall, Datenbestand etc. einzugruppieren und diese dann der entsprechend am häufigsten vorkommende Ausprägung der zu untersuchenden Zielgrösse oder ab einer bestimmten relativen Häufigkeit der jeweiligen Kategorie zuzuordnen. Im Einzelfall nehmen wir einen Fehler in Kauf. Je häufiger diese Entscheidung zu

---

treffen ist, desto eher wird der Gesamtfehler durch die so mathematisch vorgenommene Klassifikation minimiert [63]. Vorauszusetzen ist dabei, dass der einzugehende Fehler sich im Laufe der Zuordnung nicht verändert und die Parameter wie der Mittelwert erwartungstreu sind.

Für die Klassifikation stehen außer den Neuronalen Netzen weitere Verfahren zur Klassifizierung der Objekte zur Verfügung. Verfahren, die nur eine Eingruppierung vornehmen ohne Beachtung einer Zielgröße wie die Clusteranalyse oder Kohonenkarte kommen für die weitere Auswertung nicht in Betracht. Exemplarisch werden die Logistische Regression und die lineare Diskriminanzanalyse vorgestellt.

## 7.2 Logistische Regression

Die *Logistische Regression* wurde 1990 entwickelt, um die aus einem Verfahren sich ergebende Werte für die Bestimmung einer nominal skalierten Zielgröße  $y$  deuten zu können [2].

Ihre ursprüngliche Anwendung liegt im Bereich der Medizin. Es geht bei der Anwendung dieses Verfahrens primär darum, dem Patienten ein Risiko für eine bestimmte Erkrankung zuzuordnen, und erst sekundär um die Bewertung der einzelnen Risikofaktoren.

Um nominal skalierte Zielgrößen zu verarbeiten, ist die Zielgröße so zu transformieren, dass für die Zuordnung jeweils nur zwei Ausprägungen der Zielgröße betrachtet werden ("1": "stimmt" / "0": "stimmt nicht").

### Mathematische Beschreibung des Vorgehens

Um eine Zielgröße mit nur zwei Ausprägungen zu bestimmen, ist aus dem Ansatz der Linearen die Logistische Regression zu entwickeln:

$$y^j = w_0 + w_1 \cdot x_1^j + \dots + w_k \cdot x_k^j \quad (7.1)$$

Die Lineare Regression hat noch den Wertebereich aller reellen Zahlen. Um den ermittelten Wert als Wahrscheinlichkeit interpretieren zu können, ist der Bereich auf Werte zwischen 0 und 1 einzuschränken. Der geringste Wert für eine Wahrscheinlichkeit ist 0 und höchstens 1.

Anstatt nun direkt die beobachtete Gruppenzugehörigkeit zu bestimmen, wird die Wahrscheinlichkeit für die Gruppenzugehörigkeit  $P(\{\text{Gruppenzugehörigkeit}\})$  genommen.

Durch die Betrachtung der Wahrscheinlichkeit werden die Werte des Verfahrens, die zwischen den Ausprägungen "nicht-krank" und "krank" liegen, direkt als Wahrscheinlichkeit für die Zugehörigkeit deutbar.

Da bei diese Betrachtung nur eine von beiden Ausprägungen vom  $j$ -ten Fall angenommen werden kann, reicht die Betrachtung der Zugehörigkeit zu einer der beiden Ausprägungen der Zielgröße  $y$  aus. Die Einschränkung des Wertebereichs wird erreicht, indem  $y^j$  durch das logarithmierte Verhältnis der Wahrscheinlich-

---

keit für das Eintreffen des Ereignisses “Der  $j$ -te Fall gehört zur Gruppe der Kranken” zur Gegenwahrscheinlichkeit “Der  $j$ -te Fall gehört zur Gruppe mit Ausprägung “nicht-krank” ersetzt wird:

$$\ln\left(\frac{P(\{y^j = \text{“Krank”}\})}{P(\{y^j = \text{“Nicht-Krank”}\})}\right) = \ln\left(\frac{P(\{y^j = \text{“Krank”}\})}{1 - P(\{y^j = \text{“Krank”}\})}\right) \quad (7.2)$$

$$= w_0 + w_1 \cdot x_1^j + \dots + w_k \cdot x_k^j$$

Da nicht das Verhältnis, sondern die Wahrscheinlichkeit für die Zugehörigkeit zur Gruppe mit Ausprägung “Krank” von Interesse ist, wird die Gleichung nach der konkreten Wahrscheinlichkeit  $P(\{y^j = 1\})$  umgestellt:

$$P(\{y^j = \text{“Krank”}\}) = \frac{1}{1 + e^{-(w_0 + w_1 \cdot x_1^j + \dots + w_k \cdot x_k^j)}} \quad (7.3)$$

Für den  $j$ -ten Fall ergibt sich somit die Wahrscheinlichkeit zur Gruppe der “Kranken” zugeordnet zu werden. Mit der Wahl der Funktion wird angenommen, dass die Wahrscheinlichkeit einen logistischen Verlauf annimmt.

Das heißt für den Anwender, dass die Wahrscheinlichkeit für die Gruppenzugehörigkeit anfangs mit Zunahme der Ausprägung der erklärenden Variable schwach zunimmt, dann im mittleren Bereich stärker und im oberen Bereich schwächer wird: Nimmt man an, die erklärende Variable  $x_i$  hat einen positiven Einfluss auf die Zugehörigkeit zur Gruppe mit Ausprägung “krank”, so wird eine Veränderung von einem niedrigem Niveau um eine Einheit der erklärenden Variable  $x_i$  bei einem Fall nicht unbedingt dazu führen, dass der betreffende Fall der Gruppe mit Ausprägung “krank” zuzuordnen ist. Im mittleren Bereich der logistischen Funktion, in welchem ein entsprechender Fall bereits nahe an der 50%- Schranke liegt, führt die gleiche numerische Veränderung zu der Zuordnung zur Gruppe der “Kranken”.

Die Bewertung des Anstiegs einer Variablen um eine bestimmte Einheit ist in Form der Koeffizienten fest. Zwar wirkt sich die Veränderung der erklärenden Variablen  $x_i$  in der Realität je nach der Ausgangslage von  $x_i$  unterschiedlich aus, doch die Chance zur Gruppe mit der Ausprägung “Krank” zugeordnet zu werden, ändert sich bei jeder Erhöhung der Variablen  $x_i$  um eine Einheit immer um denselben Faktor, nämlich  $e^{w_i}$ . Beträgt die Chance für  $x_i = 0$  “krank”: “nicht krank” 3 : 1 und für  $x_i = 1$  die Chance 2 : 3, so hat sich die Chance “krank” zu werden um den Faktor  $e^{w_i} = \frac{2}{3} = 0.667$  verringert. Multipliziert man die Anzahl der Kranken bei  $x_i = 0$  mit dem Faktor  $e^{-1.505} = 0.222$ , so ergibt sich das Verhältnis für  $x_i = 1$ . Ist die Situation für  $x_i = 2$  zu betrachten, so wird die Anzahl an “Kranken” bei  $x_i = 1$  verwendet und mit dem gleichen Faktor wie von  $x_i = 0$  zu  $x_i = 1$  multipliziert.

Die Bestimmung der Gewichtungen bzw. der Regressionskoeffizienten für die erklärenden Variablen erfolgt ohne Betrachtung der Wechselwirkungen. Vorausgesetzt, es liegen sowohl voneinander unabhängige erklärende Variablen, als auch dem entsprechende Beobachtungen vor, erfolgt dies über eine Schätzung,

---

bei der die Maximum-Likelihood-Methode [79] verwendet wird.  
Die Vernachlässigung<sup>1</sup> der Unabhängigkeit könnte die Ergebnisse entsprechend verzerren und zu Fehlinterpretationen führen.  
Die hierfür verwendete Likelihood-Funktion ergibt sich wie folgt:

$$L = \prod_{\{y^j=1\}} P(\{y^j = \text{"Krank"}\}) \cdot \prod_{\{y^j=0\}} (1 - P(\{y^j = \text{"Krank"}\})) \quad (7.4)$$

Wenn die Gruppe derjenigen mit der Ausprägung "krank" vollständig von der mit der Ausprägung "nicht-krank" getrennt werden kann, so ist die Wahrscheinlichkeit für einen Fall mit der Ausprägung "krank"  $P(\{y^j = \text{"krank"}\}) = 1$  bzw. 100%. Vice versa für einen Fall mit der Ausprägung "nicht-krank"  $P(\{y^j = \text{"krank"}\}) = 0$  bzw. 0%.

Somit ist die Wahrscheinlichkeit für das gegenteilige Ereignis "Nicht-Krank"  $(1 - P(\{y^j = 1\})) = 1$ .

Wenn das obige Produkt den maximalen Wert von 1 erreicht, liegt dann die Gruppe derjenigen mit der Ausprägung "krank" von denen mit der Ausprägung "nicht-krank" vollständig voneinander getrennt vor.

Zur Vereinfachung der Berechnung kann die obige Funktion logarithmiert werden, womit eine Summe anstatt eines Produkts betrachtet wird:

$$\ln(L) = \sum_{\{y^j=1\}} \ln(P(\{y^j = \text{"Krank"}\})) + \sum_{y^j=0} \ln(1 - P(\{y^j = \text{"Krank"}\})) \quad (7.5)$$

Der Wert für die optimale Trennung der beiden Gruppen beträgt bei der logarithmierten Version nun  $\ln(P(\{y^j = \text{"krank"}\})) + \ln(1 - P(\{y^j = 1\})) = \ln(1) + \ln(1) = 0 + 0 = 0$ .

Bei beiden Betrachtungen sind die Extremstellen der Funktion gesucht. Um das Maximum oder Minimum einer Funktion zu bestimmen, wird die erste Ableitung der Funktion verwendet. Man betrachtet die Steigung der logarithmierten Likelihood-Funktion bei einem bestimmten Wert für den  $w_i$ - Koeffizienten. Beträgt die Steigung bei diesem Wert für den  $w_i$ -Koeffizienten 0, so liegt ein Extrempunkt vor. Da es ein Minimum in der betrachteten Funktion  $L$  nicht gibt, kann es sich nur um ein Maximum handeln.

$$\frac{\partial \ln(L)}{\partial w_i} = 0 \quad (7.6)$$

Wenn bei der Likelihood-Funktion für  $n$  Variablen  $n+1$  Koeffizienten, nämlich zusätzlich die Konstante, zu bestimmen sind, erhält man ein nichtlineares Gleichungssystem von  $n + 1$  Unbekannten und  $n + 1$  zu erfüllenden Gleichungen. Durch ein numerisches Verfahren wie das "Newton-Raphson-Verfahren" wird eine Näherungslösung ermittelt. Ein alternativer Ansatz zur Logistischen Regression ist anstatt der Verwendung des "Logarithmus" bzw.  $\ln$  der des "Sinus" bzw.

---

<sup>1</sup> S.118, [70]

---

sin. Vorteil des  $\sin$  ist, eine mögliche Periodizität in der Bewertung mitaufzunehmen. So können auch die sonst als Ausreißer erscheinenden Formen, wenn sie gleichermaßen zuzuordnen sind, mitberücksichtigt werden.

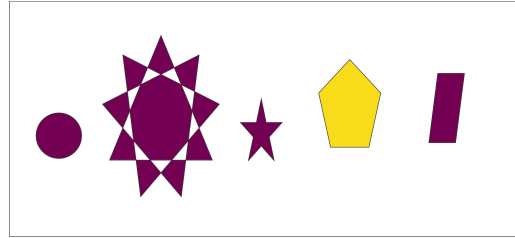


Abbildung 7.2: Differenzierung von Objektgruppen (lila vs. gelb) unter periodischer Betrachtung der Anzahl an Ecken des Objekts

**Beispiel 32.** Für einen Aktienkurs wollen wir die jeweilige Wahrscheinlichkeit bestimmen, dass der Kurs am nächsten Tag ansteigt. Dies ist ein veränderter Ansatz der bereits von uns in [46] vorgestellten Fragestellung. Anstatt den Wert zu bestimmen, wollen wir lediglich wissen, ob von einem Kursanstieg auszugehen ist. Für das  $CO_2$ -Zertifikat liegen vom 8.1.2010 bis 9.7.2010 die jeweiligen Schlusskurse des Tages vor. Die Logistische Regression setzt an allen 129 Tagen auf fallende Kurse. Für den einfachen Ansatz mit  $y = \sin(w \cdot x)$  und  $w = 7.50026$  erhalten wir 61 korrekte Zuordnungen. 15mal wurden korrekt steigende Werte vorhergesagt, 46 mal fallende. Wird der Wertebereich der Funktion eingeschränkt auf das Intervall  $[0, 1]$  durch Multiplikation mit  $\frac{1}{2}$  und Verschiebung nach oben um  $\frac{1}{2}$ , so ergeben sich 60 korrekte Zuordnungen.

### 7.3 Diskriminanzanalyse

Die *Diskriminanzanalyse* [19] ermittelt eine lineare Trennfunktion  $g$  zwischen den Gruppen mit unterschiedlicher Ausprägung der Zielgröße  $y$  anhand aller  $n$  erklärenden Variablen gleichzeitig und unabhängig von der Fallzahl bei einer bestimmten Ausprägung. Mit dem Wissen von sämtlichen Fällen und Variablen werden alle Gewichtungen  $w_1 \dots w_n$  der einzelnen Variablen geschätzt.

Die Ausprägungen der erklärenden Variablen  $x_1 \dots x_n$  werden mit den entsprechenden Gewichtungen multipliziert. Dabei ist  $w_i$  als die Gewichtung der einzelnen erklärenden Variable  $x_i$  zu verstehen. Die derart gewichteten Merkmale werden dann aufsummiert. Je nach dem wie groß der aufsummierte Wert ist, wird bei Unterschreiten einer festen Grenze der zu bestimmende Fall der einen und bei Unterschreiten der anderen Ausprägung der Zielgröße  $y$  bei jeweils zwei Kategorien zugeordnet.

#### Mathematische Beschreibung des Vorgehens

---

Wird bei der Bestimmung der Gewichte  $w_i$  der einzelnen Variablen das Verfahren nach Fisher [19] verwendet, dann berechnet sich der Wert für  $w_i$  durch den Vergleich der Unterschiede zwischen und innerhalb der Gruppen mit unterschiedlicher Ausprägung der Zielgröße  $y$ .

$w_i$  ist dann so zu wählen, dass der Unterschied innerhalb der Gruppe mit gleicher Ausprägung für die Zielgröße  $y$  bei der erklärenden Variablen  $x_i$  so gering und der Unterschied zwischen den einzelnen Gruppen mit unterschiedlicher Ausprägung der Zielgröße  $y$  so groß, wie möglich, ist.

Bei der Bestimmung des einzelnen Gewichts  $w_i$  für  $x_i$  wird in der Regel der Einfluss anderer erklärender Variablen auf  $x_i$  durch  $x_j$  und Wechselwirkungen als zumindest vernachlässigbar angesehen.

Bspw. in Form des Produktes  $x_i \cdot x_j$  könnten Wechselwirkungen explizit eingefügt werden. Da die entsprechenden Wechselwirkungen meist nicht bekannt sind und entsprechend eingefügt werden müssen, wird meist bei der Diskriminanzanalyse von der Unabhängigkeit der Variablen ausgegangen.

Die Unabhängigkeit und damit die Nichtbeachtung von Wechselwirkungen der erklärenden Variablen wird an der unten verwendeten linearen Trennfunktion  $g(x_1 \dots x_n)$  schon deutlich, die jede erklärende Variable einzeln gewichtet und keinerlei Kombination und damit Wechselwirkungen zwischen den erklärenden Variablen beinhaltet:

$$g(x_1 \dots x_n) = w_1 \cdot x_1 + \dots + w_n \cdot x_n + w_0. \quad (7.7)$$

Über die lineare Funktion  $g$  gelangt man zu der eigentlichen Zuordnung  $f$  für die Zielgröße  $y$  und damit zu der entsprechenden Ausprägung des  $j$ -ten Falles über die noch zu bestimmende Grenze oder den kritischen Wert  $K$  wie folgt.

Für die Überschaubarkeit wird das Verfahren nur bei zwei Ausprägungen beschrieben. Es ist durchaus für mehr als zwei Ausprägungen anwendbar:

$$y = \begin{cases} \text{"krank"} & \text{wenn } g(x) > K \\ \text{"nicht krank"} & \text{wenn } g(x) < K. \end{cases} \quad (7.8)$$

Liegt der durch  $g(x_1^j \dots x_n^j)$  bestimmte Wert für den  $j$ -ten Fall über dem kritischen Wert  $K$ , so ist er der Ausprägung der Zielgröße "krank" zuzuordnen, liegt er darunter, so wird er als "nicht krank" klassifiziert.

Der kritische Wert  $K$  ergibt sich als Mittel zwischen den beiden Gruppenmittelwerten für die Werte von  $g(x_1 \dots x_n)$  bei nur zwei Ausprägungen der Zielgröße.

Durchschnittlicher Wert für die Ausprägung "krank" der Zielgröße  $y$  mit der Anzahl von  $n_1$  "Krank" :

$$\overline{g(x)}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} g(x)_i \quad (7.9)$$

Durchschnittlicher Wert für die Ausprägung "nicht-krank" der Zielgröße  $y$  mit



---

der Anzahl von  $n_2$  "Nicht-Krank" :

$$\overline{g(x)}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} g(x)_i \quad (7.10)$$

Berechnung des kritischen Wertes als Mittelwert der beiden oberen Gruppenmittelwerte:

$$K = \frac{1}{2}(\overline{g(x)}_1 + \overline{g(x)}_2) \quad (7.11)$$

Die Bestimmung des kritischen Wertes als Mittelwert setzt in diesem Fall eine vergleichbare Streuung innerhalb der beiden Gruppen um den jeweiligen Mittelwert voraus. Ist die eine Streuung der beiden Verteilungen für  $g(x_1 \dots x_n)$  (getrennt nach den Ausprägungen der Zielgröße  $y$ ), die sich aus den gewichteten Verteilungen der erklärenden Variablen  $x_1 \dots x_n$  ergibt, breiter oder gibt es sogar mehrere Gipfel bei gleicher Ausprägung der Zielgröße, ist die Anwendbarkeit der Diskriminanzanalyse in Frage zu stellen.

Für die Anwendung des Verfahrens wird eine ähnliche Verteilung der Werte von  $g(x)$  für die einzelnen Ausprägungen der Zielgröße  $y$  vorausgesetzt, sowie ein Gipfel der Verteilung, der nach links und rechts abflacht.

Nur unter den genannten Bedingungen an die Verteilung kann eine optimale Trennung für die Ausprägung der Zielgröße erfolgen. Ist die Verteilung der Gruppen ungleich, so wird die Anzahl an Fehlzuordnungen sich nicht auf die Ausprägungen der Zielgröße in gleichem Umfang aufteilen. Das bedeutet, dass bei einer der Ausprägungen ein erhöhter Anteil an Fehlzuordnungen zu erwarten ist.

Um die Gewichtung in  $g(x_1 \dots x_n)$  zur Trennung der beiden Gruppen nach Fisher im Weiteren zu bestimmen, wird der folgende Ansatz gewählt: Im Zähler befindet sich die Streuung zwischen den Gruppen, im Nenner die Streuung innerhalb der Gruppen.

$$\Gamma = \frac{\sum_{i=1}^{n_{gruppen}} n_i \cdot (\overline{g(x)}_i - \overline{g(x)})^2}{\sum_{i=1}^{n_{gruppen}} \sum_{j=1}^{n_i} (g(x)_j^{n_i} - \overline{g(x)}_i)^2} \quad (7.12)$$

Wenn der Wert  $\Gamma$  maximal ist, dann ist die Streuung innerhalb der Gruppe minimal und zwischen den zu unterscheidenden Gruppen maximal.

Wenn die erklärende Variablen  $x_1 \dots x_n$  zur Unterscheidung der Gruppen in Form der Ausprägung der Zielgröße  $y$  beitragen, so sind die Ausprägungen der erklärenden Variablen innerhalb der einen Gruppe sehr ähnlich, die zwischen den jeweiligen Gruppen sehr verschieden. Um den gewünschten Unterschied zu maximieren, wird die obige Gleichung nach dem Vektor der Gewichte  $w$  abgeleitet. Es ergibt sich ein lineares Gleichungssystem aus  $n+1$  Ungleichungen und  $n+1$  unbekannten Gewichtungen, da die Konstante  $w_0$  ebenfalls zu bestimmen ist.

Da das gesuchte Maximum, der höchste Punkt der Funktion  $\Gamma$  abhängig von

---

den Gewichten  $w_1 \dots w_n$ , bei einer Steigung von 0 der Funktion erreicht wird, lautet die zu lösende Gleichung:

$$\frac{\partial \Gamma}{\partial w} = 0 \quad (7.13)$$

Das endgültige Ziel der Diskriminanzanalyse ist es, anhand der erklärenden Variablen die Gruppen mit unterschiedlicher Ausprägung der Zielgröße  $y$  linear zu trennen. Bei der Anwendung des Verfahrens wird von vornherein von einer linearen Separierbarkeit ausgegangen. Andere nichtlineare Funktionen könnten durchaus zu einer klareren Trennung der Gruppen unterschiedlicher Ausprägung der Zielgröße  $y$  führen.

### Erkenntnisgewinn durch das Verfahren

Die ermittelte Zuordnung zu den Ausprägungen der Zielgröße  $y$  durch die erklärende Variablen  $x_1 \dots x_n$  ist anhand der Gewichtungen leicht nachvollziehbar. Der Betrag der Gewichte gibt die Stärke des Einflusses an. Ohne den Betrag erhält man durch das Vorzeichen des Gewichts die Richtung des Einflusses. Dem zufolge bedeutet ein großes Gewicht im Betrag einen erheblichen Einfluss auf die Zielgröße.

Werden die Ausprägung beim kritischen Wert in der Form "nicht-krank" < "krank" der Zielgröße  $y$  angeordnet, so bedeutet ein negatives Gewicht  $w_i$ , dass mit dem Einfluss von  $x_i$  die Zugehörigkeit zur Ausprägung "nicht krank" geringer wird, während bei positiven Gewicht  $w_i$  mit steigenden Einfluss von  $x_i$  die Zugehörigkeit zur Ausprägung "krank" der Zielgröße  $y$  wächst. Mit jeder Erhöhung um eine Einheit der erklärenden Variablen  $x_i$  steigt der Wert von  $g(x_1 \dots x_n)$  konstant an.

### Allgemeine Vor- und Nachteile

Ein erheblicher Vorteil für den Anwender ist, dass die lineare Trennfunktion und die sich ergebende Zuordnung leicht ausführbar und die Einflüsse aus der Trennfunktion ohne weiteres ablesbar sind.

Nachteilig ist, dass bestimmte Verteilungen ungeprüft angenommen werden. Nur aufgrund der Verteilungsannahmen wird eine annähernd gleiche Fehlklassifikation bei den Ausprägungen der Zielgröße vermutet.

In die Trennfunktion geht die Annahme der Linearität ein. Bei der erklärenden Variable "Distanz in der Frontalansicht" wird somit a priori angenommen, dass mit jeder Verringerung die Zugehörigkeit zu einer der Gruppen der Zielgröße "Diagnose" konstant zunimmt oder abnimmt. Des weiteren werden die erklärenden Variablen  $x_1 \dots x_n$  als metrisch oder quasimetrisch bewertet.

Liegen erklärende Variable nominaler Skalierung vor, so sind diese in eine oder

mehrere Variablen quasimetrischer, nämlich binärer Skalierung erst zu transformieren.

Ordinal skalierte Merkmale sind entweder ebenfalls in binäre Skalierung zu überführen oder als metrisch zu betrachten. Das heißt, dass der Abstand zwischen aufeinander folgenden Ausprägungen als gleich angenommen wird.

Die Werte der linearen Diskriminanzfunktion können nur in bezug auf den kritischen Wert  $K$  gedeutet werden. Für den Vergleich der einzelnen Patienten untereinander eignen sie sich nur unter Bezug auf den kritischen Wert  $K$  und der damit verbundenen Zuordnung zu den Kategorien der Zielgröße  $y$ .

Ein erweiterter Ansatz der Diskriminanzanalyse für die nichtlineare Trennung kann

$$g(x_1 \dots x_n) = \ln(w_1 \cdot x_1 + \dots + w_n \cdot x_n + w_0) \quad (7.14)$$

sein.

Im unteren Bild wird veranschaulicht, dass der nichtlineare Ansatz eher in der Lage ist, die gelben von den grünen Objekten zu trennen.

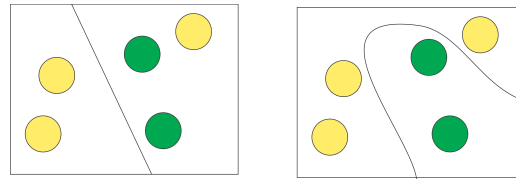


Abbildung 7.3: Differenzierung der Objektgruppe: Linear versus Nichtlinear

## 7.4 Chaotisches Verhalten der verwendeten Verfahren

Bei der Bestimmung und Verwendung einer Zuordnung zur Bestimmung einer Zielgröße stellt sich die Frage, ob chaotisches Verhalten bei der Bestimmung der Zuordnung zu erwarten ist. Das heisst, ob kleine Veränderungen bereits zu grösseren Veränderungen führen. Gerade bei Neuronalen Netzen wird häufig davon ausgegangen, dass sie mit Messfehlern oder Rauschen besser umgehen können als andere Verfahren. Doch gerade der Iterationsprozess zur Bestimmung der Zuordnung in Neuronalen Netzen kann chaotisches Verhalten aufweisen. Dies kann auch bei entsprechenden Voraussetzungen bei der Logistischen Regression eintreten.

Um chaotisches Verhalten nachzuweisen, ist zunächst die Ergodizität zu prüfen. Der entsprechende Nachweis für die Sigmoidfunktion ist in [58] erbracht worden. Die Sigmoidfunktion wird sowohl in der Logistischen Regression als auch in den Neuronalen Netzen verwendet. Darüber hinaus ergibt sich meist durch die Wahl

---

der Eingabe, dass die eingegebenen Werte wieder auf sich selbst abgebildet werden, womit sicher gestellt wird, dass der gesamte Wertebereich abgedeckt wird. Sowohl bei den Neuronalen Netzen als auch bei der Logistischen Regression handelt es sich also um eine Selbstabbildung, wenn die Eingangswerte und die Ausgangswerte im Intervall zwischen 0 und 1 liegen.

Im Weiteren ist, wenn die Ergodizität und Selbstabbildung gegeben ist, der Lyapunov-Exponent zum Nachweis des chaotischen Verhaltens zu berechnen<sup>2</sup>.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \left| \frac{df^n}{dx}(x) \right| = \lambda_f(x) > 0 \quad (7.15)$$

$$f(x) = \frac{1}{1 + e^{-w \cdot x}} \quad (7.16)$$

Für die  $n$ -te Iteration bedeutet dies:

$$f^n(x) = \frac{1}{\frac{-w}{\frac{-w}{\frac{-w}{1 + e^{1+e^{1+e^{1+e^{\dots}}}}}}}} \quad (7.17)$$

Differenziert nach  $x$  ergibt sich:

$$\frac{df^n(x)}{dx} = w^n \cdot R \quad (7.18)$$

$$R = \frac{e^{-w \cdot \frac{1}{1+e^{\dots}}}}{(1 + e^{\frac{w}{1+e^{\dots}}}) \dots (1 + e^{-w} 1 + e^{-w \cdot x})^2 (1 + e^{-w \cdot x})^2} \quad (7.19)$$

Für die Grenzwertbetrachtung ergibt sich:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln |w^n R| = \lim_{n \rightarrow \infty} \frac{1}{n} (n \ln(w) + \ln |R|) \quad (7.20)$$

Da  $|R| > 0$ ,  $R \neq 1$ , der Logarithmus dieses Wertes multipliziert mit  $\frac{1}{n}$  gegen Null geht, sowie  $\ln w$  nur dann 0 ist, wenn  $w = 1$ , gilt für  $\lambda_f(x) > 0$ . Damit kann chaotisches Verhalten der Sigmoidfunktion bei unter den entsprechenden Bedingungen vorliegen.

Nach [58] liegt stochastische Stabilität vor. Die stochastische Stabilität sagt jedoch nichts über die Stabilität der Gewichtungen aus. Gerade in den Anwendungen wird die Ausgabe des Netzes in Abhängigkeit von der Eingabe interpretiert. Dies führt zu unklaren oder gar konträren Aussagen über die Bewertung der Eingabe im Netz.

**Beispiel 33.** In der von mir durchgeführten Rückfalluntersuchung in Rheinland-Pfalz für die Jahrgänge 1996 bis 2000 [22] wird das Beispiel für die Disziplinarmaßnahmen bei Verstößen gegen die Hausordnung im Strafvollzug gewählt.

---

<sup>2</sup> vgl. S.73, [56]

---

Es ergibt sich sogar eine Rückfallreduzierung mit dem Erhalt einer Disziplinarmaßnahme bei einem der Durchgänge. In dem berechneten Modell würde das bedeuten, dass Jugendliche, die gegen die Regeln im Vollzug verstossen, weniger rückfällig werden. Zu dieser Aussage gelangt man nur, weil die Anfangsgewichte zufällig und nur geringfügig anders gewählt werden.

## 7.5 P-KFA

Die Prädiktions-Konfigurationsfrequenzanalyse [75] ermittelt die Relevanz von Merkmalkombinationen für die Zuordnung zu einer Kategorie. Sie betrachtet nicht ein erklärendes Merkmal allein, sondern die Relevanz der Merkmalskombination. Wenn Merkmalskombinationen häufiger vorkommen als in Folge des Zufalls zu erwarten ist, wird von einem Typ gesprochen. Ist eine Merkmalskombination seltener als der Zufall es erwarten lässt, handelt es sich um einen Antityp. Der für die Approximation der Normalverteilung verwendete  $z$ -Wert erreicht beim Typ entsprechende positive numerische Werte und beim Antityp negative. Die Betrachtung von Merkmalskombinationen erlaubt es, die bekannten Wechselwirkungen zwischen den Merkmalen in Form einer Kombination zu betrachten. Die Skalierung der Merkmale ist in diskrete Stufen zu betrachten. Stetige Merkmale ohne entsprechende Vorverarbeitung können nicht mit der KFA untersucht werden.

Da in den Sozialwissenschaften, in der Psychologie und insbesondere auch in der Medizin Daten von unterschiedlichem Meßniveau vorliegen, ist ein Verfahren, welches das nominale Meßniveau und damit die geringsten Voraussetzungen an die Skalierung stellt, bevorzugt zu wählen.

Die Prüfung des Zufalls erfolgt anhand der in den erhobenen Daten vorliegenden hypergeometrischen Verteilung oder bei entsprechender Begründung anhand der Näherung durch die Normalverteilung ( $z$ -Wert). Der Grundgedanke der Konfigurationsfrequenzanalyse wurde von Lienert [68] dargestellt.

Die Prädiktions-Konfigurationsfrequenzanalyse P-KFA baut auf der KFA auf. Ihr Ziel ist es, die für die Erklärung eines anderen Merkmals relevanten Merkmalskombinationen zu bestimmen. Das bedeutet, dass bei der P-KFA die signifikanten Merkmalskombinationen ermittelt werden, die eine Zuordnung zu einer der Kategorien des zu untersuchenden Merkmals oder der Zielgröße ermöglichen. Aufgrund des Ansatzes eines häufigeren Auftretens einer Merkmalskombination als in Folge des Zufalls zu erwarten wäre, ermittelt die P-KFA nur Typen und keine Antitypen. Vorteil des Verfahrens ist es, dass klar beschreibbare Typen ermittelt werden und nicht nur die Aussage über ein einzelnes erklärendes Merkmal getroffen werden. Der Ansatz der P-KFA wurde bereits in der von uns vorgenommene Rückfalluntersuchung in Rheinland-Pfalz [24] im Bereich der Kriminologie verwendet.

**Beispiel 34.** In dem von uns erstmalig in Italien publizierten Ansatz für die Rückfalluntersuchung [24] wird die Relevanz des "Sozialen Verlierers", Personen ohne Schul- und Berufsabschluss, sowie Heimwechsel in der Zuordnung zur "Rückfälligkeit" in Form eines erneuten Eintrags deutlich. Sie werden häufiger

---

rückfällig als vom Zufall her zu erwarten wäre. Entsprechend kann der Jugendstrafvollzug in den jeweiligen Bundesländern in seiner Behandlung entsprechend klientelorientiert tätig werden [48].

## Kapitel 8

# Anwendungen der Formanalyse

Für die bisherige Anwendung der Formanalyse sind mindestens zweidimensionale Daten vorauszusetzen. Bei der Bestimmung einer "mittleren Form" wird im Weiteren davon ausgegangen, dass es eine entsprechende Verteilung gibt, die eine eindeutige mittlere Form zulässt, sowie dass diese mittlere Form für die Anwendung zur Klassifizierung der Objekte von Bedeutung ist.

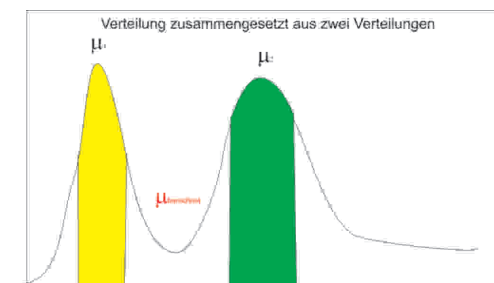


Abbildung 8.1: Differenzierung der Objektgruppen anhand der Verteilung

**Beispiel 35.** Bei der Untersuchung von Affenschädeln in Mardia und Dryden [69] ist von einem zu beachtenden Unterschied zwischen männlichen und weiblichen Schädeln auszugehen. Eine mittlere Form ist nur unter Beachtung des Geschlechts von Bedeutung. Sollte eine generelle mittlere Form der Affenschädel bestimmt werden, so wäre dem jeweiligen Anteil der Geschlechter Rechnung zu tragen

Bei aller Mathematik sind somit auch die Probleme des Anwenders und die bereits vorhandenen Kenntnisse zu sehen.

---

Im Weiteren werden die drei in dieser Arbeit vorliegenden Anwendungsgebiete in der Medizin, im Ingenieurwesen und in der Kriminologie im Detail vorgestellt.

## 8.1 Nierentumore im Kindesalter

Um eine medizinische Diagnose stellen zu können, greift der Arzt u.a. auf Bildaufnahmen in Form der Computertomographie, der Magnetresonanztomographie, sowie Ultraschall und Röntgen zurück. Trotz des medizinischen Wissens und der Erfahrung mit anderen Patienten gelingt es dem Arzt nicht in allen Bereichen eine sichere Diagnose bereits anhand der vorliegenden Bilder eines Patienten abzugeben. Ein differentialdiagnostisch schwieriger Bereich stellen die Nierentumore des Kindesalters dar. In diesem Fall wird die onkologische Therapie ohne histologische Sicherung eingeleitet, da eine Biopsie durchaus zu einer Tumorruptur oder zu Streuung von bösartigen Zellen führen kann. Der Radiologe hat somit eine Schlüsselstellung in der Therapieentscheidung.

Ein Großteil der Nierentumore im Kindesalter stammen aus der Gruppe der Nephroblastome, die so genannten *Wilms-Tumore* [88]. Wilms-Tumore treten meist in den ersten sechs Lebensjahren auf. Pro Jahr kann man von ca. 100 - 130 neu diagnostizierten “Wilms-Tumoren” ausgehen. Wilms-Tumore treten bei beiden Geschlechtern gleichermaßen auf. Sie stellen in Europa 6 – 7% aller bei Kindern diagnostizierten Tumore dar. In Asien tritt der Wilms-Tumor im Vergleich zu Europa seltener auf [15]. Die Unterschiede in der Häufigkeit des Wilms-Tumors sind weniger geographischer Natur, sondern hängen eher mit der Zugehörigkeit zu ethnischen Gruppen zusammen. Bei der Entstehung des Tumors geht man vor allem von genetischen Faktoren aus [12]. So macht bei der schwarzen Bevölkerung sowohl in Afrika, als auch in Nordamerika der Wilms-Tumor um die 10% [87] der bei Kindern diagnostizierten Tumore aus.

Die Gruppe der Wilms-Tumore bzw. Nephroblastome läßt sich in 3 Risikoklassen unterteilen: *I: niedrige Bösartigkeit*, *II: Standardrisikotyp* und *III: hohe Bösartigkeit*. Jede Risikoklasse wiederum in bestimmte histologisch-pathologische Typen. Die Risikoklasse II bspw. in *a: epithelialer Typ*, *b: stromareicher Typ*, *c: Mischtyp* und *d: regressiver Typ*. Die Risikoklasse III weist u.a. den *a: blastemreichen Typ* auf.

*Neuroblastome* werden pro Jahr ebenfalls ca. 130-140 Fälle neu registriert [54]. Neuroblastome entwickeln sich aus embryonalen Nervenzellen. Um die 60% der Primärmanifestationen liegen im Abdomen und sind von den Wilms-Tumoren zu unterscheiden.

*Nierenzellkarzinome* treten meist bei Erwachsenen auf und sind im Kindesalter sehr selten. Sie machen ca. 3 – 6% der Nierentumore im Kindesalter aus [16]. Vom Bild her können sie bislang nicht sicher in jedem Fall vom Wilms-Tumor unterschieden werden. Da bis zum 16.Lebensjahr die Wilms-Tumore im Vergleich zu den Nierenzellkarzinomen überwiegen, geht der Arzt bei Patienten unter 16 Jahren meist vom Wilms-Tumor aus.

*Klärzellensarkome* sind Tumore mit hohem Malignitätsgrad. Typisch sind Absiedlungen in den Knochen, der Lunge und dem Gehirn. Sie kommen ca. genauso



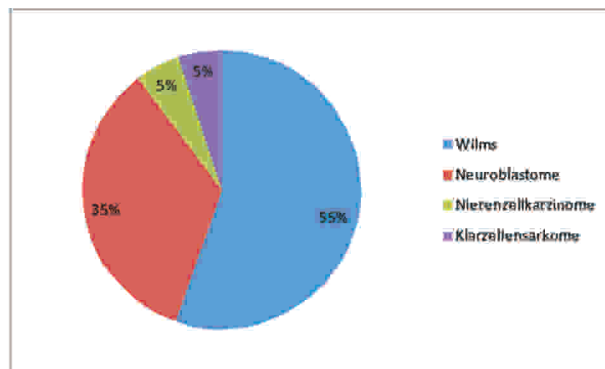
häufig vor wie die Nierenzellkarzinome.

Hervorzuheben ist hinsichtlich der Suche nach einer zu erwartenden mittleren Form von Neuroblastomen, dass diese im Gegensatz zu Wilms-Tumoren, Klarzellensarkomen und Nierenzellkarzinomen keine klar vorgegebene Lage im Körper haben, somit vermutlich in ihrer Form und Gestalt weniger an Gegebenheiten des Körpers gebunden sein könnten und von daher vermutlich eine höhere Varianz aufweisen.

Tabelle 8.1: Auftreten der Tumorarten

Tumorart	Häufigkeit in %
Wilms- Tumor	55.3%
	$(h_{Wilms} = 100 - 130)$
Neuroblastome	34.4%
	$(h_{Neuroblastom} = 81)$
Nierenzellkarzinome	5.1%
	$(h_{Nierenzellkarzinom} = 12)$
Klarzellensarkome	5.1%
	$(h_{Klarzellensarkome} = 12)$

*Anmerkung: Angaben für Neuroblastome und Wilms-Tumore als Näherungen aus [www.kinderkrebsinfo.de](http://www.kinderkrebsinfo.de), für Nierenzellkarzinome und Klarzellensarkome u.a. aus dem Urologen Nr. 42 Bd. 2*



Von der Zuordnung des Patienten mit dem vorliegenden Tumor zu einer Tumorart wie in Tabelle 8.1 durch den Arzt hängt das weitere Vorgehen und damit die Gesundheit des Patienten ab: Während bei einem Patienten mit Wilms-Tumor der Einsatz der Chemotherapie sinnvoll ist, kann sie bei den anderen

---

Tumorarten erst mit geänderten Therapieschema - Zusammensetzung und Intensität - von Nutzen sein oder gar sinnlos wie z.Bsp. beim Nierenzellkarzinom. Die Fehlklassifikation eines Wilms- Tumors als Nicht-Wilms-Tumor kann ohne Chemotherapie zur primären OP führen oder macht eine transkutane Biopsie inklusive deren Risiken notwendig. Die Primär- OP bei Wilms- Tumor hat eine höhere Komplikationsrate gegenüber der OP nach präoperativer Chemotherapie. Ohne chemotherapeutische Vorbehandlung sind Tumorrupturen intraoperativ häufiger, so dass nachfolgend dann eine abdominelle Strahlentherapie im Fall einer Ruptur notwendig wird. Somit ist jegliche Methode einer genauen Differenzierung der Tumorarten vor Therapiebeginn von größtem Interesse, um die optimale Therapie zu gewährleisten.

Um die Zuordnung des Patienten und des Tumors zu einer bestimmten Gruppe von Patienten bzw. zu einer Tumorart, somit die Diagnose zu verbessern, könnten mathematische Verfahren, die die Form des Tumors anhand der Bilder mehrerer Patienten untersuchen und die verschiedenen Patienten und die jeweilige Tumore miteinander vergleichen, wertvolle Dienste leisten. Zur Verfügung stehen den mathematischen Verfahren bei den Nierentumoren im Kindesalter für die Auswertung der in der Regel bei dieser Art klar erkennbare äussere Rand des jeweiligen Tumors.

---

## 8.2 Räumliche Analyse des menschlichen Verhaltens

Der Mensch macht sich zum einen den ihm zur Verfügung stehenden Raum nutzbar durch sein Verhalten und unterliegt andererseits den Einschränkungen des Raumes. Die Analyse des räumlichen Verhaltens einer Handlung sollte Aufschluss über den Handelnden geben und seine Möglichkeiten, sich den Raum nutzbar zu machen.

In der Kriminologie als Lehre des Verbrechens und in der Kriminalistik als Ermittlung geeigneter repressiver und präventiver Maßnahmen ermöglicht die Analyse des räumlichen Verhaltens zum einen Aussagen über das Verbrechen und den Verbrecher an sich und zum anderen bietet sie Ansätze zur Ergreifung des Täters und zum Schutz des Opfers.

Die absolute Zahl an Tötungsdelikten im Bundesgebiet betrug im Jahr 2006 insgesamt 818 Fälle. Davon sind um die 23 in Zusammenhang mit Sexualstraftaten zu sehen. Die Taten, die auf das Konto von "Serienmördern" gehen, sind in den 23 Fällen bereits enthalten.

Trotz der geringen Fallzahl genießen "Sexualmorde" ein erhebliches öffentliches Interesse, insbesondere wenn sie Serientätern zuzurechnen sind. Sowohl das stete öffentliche Interesse an derartigen Fällen, als auch das damit verbundene Gefühl der Unsicherheit für den Mitbürger führen zu dem Druck auf die ermittelnden Behörden, zügig den jeweiligen Täter zu fassen.

Die operative Fallanalyse ist anfangs eigens für den Zweck der zügigen Täterergreifung eingeführt worden. In die Methodik der operativen Fallanalyse halten mittlerweile, auf der Suche nach weiteren Erkenntnissen über das Verhalten und die Merkmale der Täter, auch quantitative Methoden verstärkt Einzug.

Der Einsatz quantitativer Methoden erfordert meist eine äußerst gründliche Auseinandersetzung des Kriminologen und Kriminalisten mit den mathematischen Voraussetzungen (Verteilungen, Skalierung der Merkmale etc.) und der Deutbarkeit der Ergebnisse. Der mögliche Erkenntnisgewinn quantitativer Methoden ist schon bei der Erfassung sozialwissenschaftlicher Daten stets zu hinterfragen: Berechtigen die erfassten Fälle den Kriminologen schon zu generellen Aussagen? Reicht die Anzahl an Fällen aus, um auf weitere Fälle zu schließen und sind die Fälle untereinander überhaupt vergleichbar? Im Folgenden wird die bisher hauptsächlich in der Mathematik genutzte Formanalyse zur Analyse des räumlichen Verhaltens von kriminologisch interessanten Tätergruppen auf ihre Verwendbarkeit hin untersucht. Vorteile der Formanalyse in der operativen Fallanalyse sind:

- Alle für die Begehung der Tat relevanten Orte sind gleichzeitig in die Untersuchung miteinbezogen. Es wird das gesamte räumliche Profil des Täters, welches sich in seinen Handlungsabläufen aus mehreren Orten zusammensetzt, auf Rückschlüsse zum Täter hin und seine Charakteristika geprüft.

- 
- Täter mit unterschiedlicher Mobilität bei Begehung der Tat werden vergleichbar.

Der Raum des Täters hat einen erheblichen Einfluss auf die Begehung der Tat. Je nach Raum und seiner Struktur hat der Täter die Möglichkeit, ein entsprechendes Opfer zu finden, den Ort für die Tötung zu wählen und die Leiche zu verbergen. Alle Orte der Taten strukturieren den Raum.

Unabhängig von den Möglichkeiten des Raumes, insofern sie für den Täter zur Verfügung stehen, sind aus dem räumlichen Vorgehen Rückschlüsse auf den Täter und sein Verhalten zu ziehen. Die Merkmale des Täters schlagen sich in einem für ihn charakteristischen räumlichen Verhalten nieder. Werden aufgrund des räumlichen Vorgehens Rückschlüsse auf die Merkmale des Täters gezogen, wird vorausgesetzt, dass Handlungen eines Täters durch den Lebensraum und die darin implizierten Alltagsroutinen geprägt sind. Es ist anzunehmen, dass der Täter bei einem Tötungsdelikt auf die ihm alltäglich zur Verfügung stehenden oder zumindest bekannten Ressourcen im Raum zurückgreifen wird. Dies könnte dem Täter das für die Begehung der Tat nötige Gefühl der Sicherheit geben.

Die untersuchten Fälle von Serientätern stammen aus der Studie "Geographische Verhaltensmuster bei Serien-Sexualmördern" von S. Harbort (2006). Als Serien-Sexualmörder kommt demnach nur ein sexuell motivierter Täter mit mindestens zwei voneinander unabhängig begangenen Taten in Frage. Von 1965 bis 2005 stehen 49 Täter für die Untersuchung zur Verfügung. Es geht bei dieser Untersuchung nur um die Ermittlung räumlicher Charakteristika, eine brauchbare Klassifikation ist aufgrund der geringen Fallzahl nicht zu erwarten und durchzuführen. Allein aufgrund des betrachteten Zeitraums (1965-2005) ist das Ziel der Anwendung der Formanalyse in dieser Arbeit auf die Brauchbarkeit des Verfahrens beschränkt.

## 8.3 Elektronische Nasen

Elektronische Nasen werden in der Lebensmittelindustrie eingesetzt. Neuere Anwendungen werden in Kooperation mit der Universität Kassel im Bereich Geruchsbestimmung in der Kanalisation und Detektion von Sprengstoff und Drogen durchgeführt. Bei Geruch gibt es die quantitative Maßeinheit GE/ $m^3$ . Eine Geruchseinheit entspricht dabei der Menge von Geruchsstoffen in 1  $m^3$  Neutralluft, die an der Geruchsschwelle eine Geruchswahrnehmung auslösen laut entsprechender DIN-Norm [14]. Die Einheit ist somit genormt.

Die qualitative Geruchsbelastung konnte bisher nicht genau beschrieben werden. Doch erst durch die Beschreibung des qualitativen Geruchs sind Aussagen über die Geruchssubstanz möglich. Die Formanalyse bietet einen möglichen Ansatz, nicht nur die Quantität, sondern auch die Qualität von Geruch mathematisch zu beschreiben.

Dafür sind die jeweils spezifischen Gase über Elektronische Nasen messbar. Eine Elektronische Nase besteht aus mehreren Sensoren. Jeder Sensor misst ein spezifisches Gas oder Verbindung. Die Messung erfolgt über die Änderung des

---

elektrischen Widerstands oder der Frequenz des Sensors durch die angelagerten Gase. Für die Verwendung Elektronischer Nasen sind die Sensorsignale für den Anwender hinsichtlich der Erkennung einer Geruchssubstanz zu deuten. Das Sensorsignal an sich nützt dem Anwender ohne entsprechende Deutung des einzelnen Sensors oder mehrerer Sensoren als Profil nichts.

Im Bild 8.2 ist die unserer Untersuchung zugrundeliegende Elektronische Nase abgebildet. Ihr Messprinzip beruht auf der Änderung des elektrischen Widerstands.



Abbildung 8.2: Beispiel einer Elektronischen Nase von Airsense: Pen 2

### 8.3.1 Beschreibung der Geruchsqualität durch Unterscheidung von Standorten

In unserer ersten Anwendung galt es, das Geruchsprofil aus der Kanalisation von den Geruchsprofilen im ländlichen Raum zu differenzieren. Hierfür wurden Messungen mit der Elektronischen Nasen am Emscher Kanal vorgenommen und mit den Messungen bei Schweinegülle und Biogas verglichen, wie dies in Giebel, Franke, Frechen (2008) vorgestellt wurde [27]. Die ersten vorgestellten Ergebnisse weisen schon auf eine Brauchbarkeit des Verfahrens hin, die Qualität der Gerüche in Form eines Profils aus mehreren Sensoren differenzieren zu können.

### 8.3.2 Detektion von Drogen und Sprengstoff über die Geruchsqualität

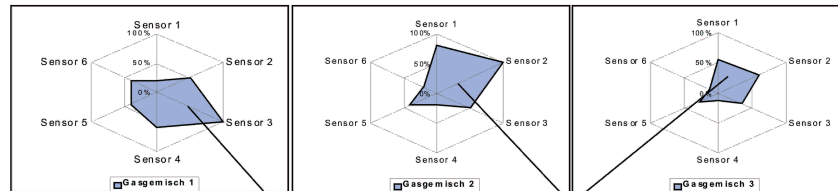
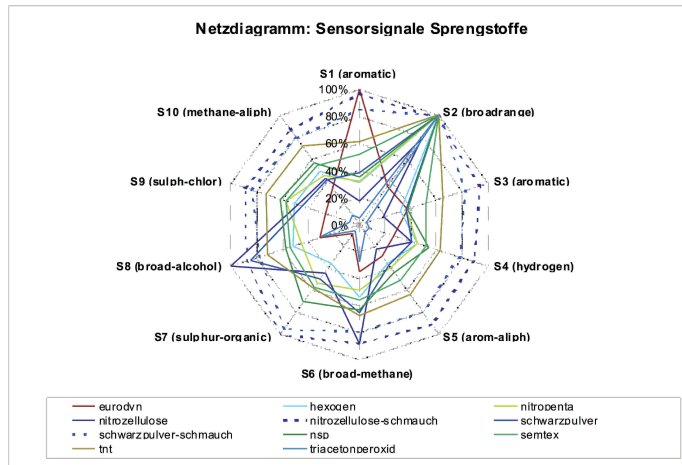
Nicht die Intensität des Geruchs, sondern die Qualität des Geruchs bietet die Möglichkeit gefährliche und illegale Substanzen zu detektieren. Polizeispürhunde können nur eine sehr begrenzte Zeit zur Suche nach Sprengstoffen und Drogen eingesetzt werden. Für die Überwachung des öffentlichen Raums wie Flughäfen, Bahnhöfe, Gerichtsgebäude, Strafanstalten, Polizeistationen und bei politischen Anlässen ist daher eine kontinuierliche Detektion von Sprengstoffen und Drogen nötig.

Allein in Nordhessen handelt es sich um mehr als 100 Einsätze zur Sprengstofferkennung pro Jahr. Eine Detektion von Sprengstoffen und Drogen ist über die Umgebungsluft und die in der Luft enthaltenen spezifischen Gase möglich. Für

die Interpretation der Sensorsignale sind erstens Messungen mit Sprengstoffen und Drogen durchzuführen und zweitens ein Verfahren zu entwickeln, dass bei weiteren Messungen zumindest die Wahrscheinlichkeit für das Vorhandensein eines Sprengstoffs oder Drogen angibt.

Für jede Geruchssubstanz ergibt sich ein zweidimensionales Profil. Für die Verwendbarkeit dieses Profils sind weitere Messungen mit den entsprechenden Störsubstanzen nötig. Einige Störsubstanzen können den Geruch völlig maskieren. Dieser Maskierungsvorgang ist durch Verfeinerung der Messtechnik noch zumindest zu erschweren. Die Formanalyse kann in diesem Fall nur zur Prüfung der Signifikanz des gemessenen Profils dienen. Bisher reicht jedoch die Anzahl an Messungen aufgrund der Vergleichbarkeit der Messung, der zu untersuchenden Störstoffe etc. noch nicht aus ( $n = 40$ ), um ein entsprechendes Detektionsverfahren zu entwickeln.

Die folgende Abbildung 8.3.2 zeigt mehrere Geruchsprofile von Sprengstoffen.



“Fingerabdruck” 1      “Fingerabdruck” 2

Abbildung 8.3: Geruchsprofile von Sprengstoffen

---

### 8.3.3 Bestimmung der Geruchsquantität bzw. Geruchsbelastung

Die Abwasserkonditionierung durch Dosierung mit Chemikalien in Abwasserkanälen dient dem aktiven Korrosionsschutz und ist eine Möglichkeit zur Begrenzung der Geruchsstoffbelastung im Abwasser. Die normalerweise verwendeten Chemikalien wirken oxidierend, das heißt, sie verhindern ein Anfaulen des Abwassers oder transformieren die Geruchsstoffe, z.B. Sulfide in eine bezüglich des Geruchs nicht mehr wahrnehmbare unlösliche Verbindung (Fällung zu Schwefeleisen  $\text{FeS}$ ), so dass Schwefelwasserstoff  $\text{H}_2\text{S}$ , einer der wichtigsten, starken Abwasser-Geruchsstoffe gebunden wird. Die Dosierung erfolgt entweder konstant oder gesteuert und basiert auf der Erfahrungen.

Steuerparameter sind z.B. der Abwasser-Durchfluss oder die  $\text{H}_2\text{S}$ -Konzentration in der Kanalluft (also in der Gasphase). Eine Bewertung der Kanalluft als Grundlage der Dosierung ist jedoch nicht immer zielführend, zum Teil ist auch das Potential im Abwasser zu berücksichtigen.

Grundlage für die wirtschaftliche Regelung einer Chemikaliendosierung zur Sulfid- und Geruchsstoffminderung im Abwasser ist eine optimierte bedarfsorientierte Abwasserbehandlung. Das bedeutet, dass eine teure Überdosierung in Zeiten geringer Sulfid- und/oder Geruchsbelastung vermieden wird, was die Kosten infolge Chemikalieneinsatz minimiert, und dass in Zeiten hoher Belastungen Chemikalien in ausreichendem Umfange eingesetzt werden, um Korrosion und Probleme infolge Geruch zu vermeiden und insoweit mögliche Folgekosten entweder aus baulichen Sanierungsnotwendigkeiten oder aus Kosten, die sich im Zusammenhang mit Maßnahmen zur Geruchsbekämpfung oder auch dem Verlust an Ansehen in der Bevölkerung ergeben.

Eine optimale Dosierung muss zwei Bedingungen erfüllen, zum einen die Verwendung von so wenig Dosiermittel wie nötig und zum anderen die Verwendung einer immer ausreichenden Dosiermenge bzw. die Wahl der geeigneten Chemikalie, um indirekte Kosten aufgrund von Beschwerden oder Instandsetzungsmaßnahmen einzusparen.

Beides ist nur mit einer prozessgesteuerten Dosierstrategie und einer quasi-online Messung von Sulfid- und Geruchsstoffkonzentration zu erreichen. Daher müssen für die Bestimmung des Geruchs anstelle von Probanden Multisensorarrays eingesetzt werden. Ein angepasstes mathematisches Modell bewertet die Rohwerte des Multisensorarrays in Bezug auf "Geruch" und schätzt die vorhandene Geruchsbelastung.

In der folgenden Abbildung ist einer der möglichen Einsatzorte dargestellt:



Abbildung 8.4: Kanalabschnitt mit Geruchsbelastung



## Kapitel 9

# Auswertung der Daten

### 9.1 Nierentumore bei Kleinkindern

#### 9.1.1 Erfassung der Daten

Von mehr als 150 CDs konnte nur ein Teil der Patienten in die Erhebung eingehen. Grundlage der Entscheidung war die Brauchbarkeit des Bildmaterials. Patienten, bei denen zu wenig Bildmaterial vorlag oder die Bildqualität einer über die Sequenz durchgängigen Erfassung des Tumors entgegenstand, mussten aus der Erhebung herausgenommen werden. Zu beachten ist, dass ca. ein Viertel aller Patienten in Deutschland keine MRT-Aufnahmen haben und damit gar nicht erst als CD vorliegen können.

Die Daten der 111 erfassten Patienten sind zunächst in eine für die Formanalyse verwendbare Form zu überführen. Allein bei der Erfassung ist darauf zu achten, dass für alle Patienten die Landmarken in vergleichbarer Weise erfasst werden können. Für 3 Patienten konnte keine Diagnose ermittelt werden. Davon war 1 Patient mittlerweile bereits verstorben. Das Alter der Patienten zum Zeitpunkt der Aufnahme des MRT liegt bei 4 Jahren und 2 Monate.

Die Gruppe der erfassten Patienten setzt sich aus 85 Nephroblastomen (78,7%) zusammen und aus 23 Nicht-Nephroblastomen (21,3%). Damit entsprechen die Anteile einem Jahrgang von zu diagnostizierten Tumoren in der Uniklinik Heidelberg. Unter den Nicht-Nephroblastomen befinden sich 9 Neuroblastome, einer davon mit zwei Knoten in der Nähe des Pankreas, 4 Klarzellensarkome, 5 Nierenzellkarzinome, 2 Rhabdoidtumore und 1 nicht näher diagnostiziertes Nicht-Nephroblastom. Gerade die mathematisch verwertbare Erfassung der Tumore erweist sich als äußerst schwierig. Die Qualität der Aufnahmen hängt vom Gerät und vom Klinikum ab.

Jeder Tumorrund wurde in Absprache mit Herrn Privatdozent Dr. JP Schenk (Referenzradiologe der Nephroblastomstudie SIOP 2001 / GPOH ) ermittelt. Trotzdem ist von einer gewissen medizinischen Unschärfe bei der Einschätzung des Tumorgewebes auszugehen. Eine exakte Messung ist anhand des Bildmaterials und aufgrund des Untersuchungsgegenstandes (Diskriminierung des Tumor-

---

randes anhand visuell wahrgenommenen Intensitätsdifferenzen zwischen Tumorgewebe, Tumorkapsel und benachbartes normales Körpergewebe) nicht möglich. Zunächst sind alle Bilder im DICOM-Format zu einem dreidimensionalen Objekt zusammenzufügen. Dies ist mit unterschiedlichen Programmen wie den hier verwendeten 3D-Doctor [1] möglich. In der Abbildung 9.1.1 ist eine solche Sequenz exemplarisch dargestellt. Daraufhin ist eine vergleichbare Erfassung der Landmarken sicher zu stellen, um die Form als wirkliches Unterscheidungskriterium der Tumore zu erhalten. Dass die Form nach Normierung und Zentrierung im Vergleich zur Gestalt nach Zentrierung etc. das geeignete Entscheidungskriterium ist, wurde bereits in der Vorabuntersuchung [38] festgestellt. Zusätzlich zu den bisher zur Herstellung der Vergleichbarkeit verwendeten Verfahren wird erstmalig die Reflexion eliminiert. Das bedeutet, dass alle Tumore nach rechts gespiegelt werden. Gerade im dreidimensionalen Bereich zeigt sich die Bedeutung der an dem Tumor angrenzenden Organe. Im zweidimensionalen Bereich konnte noch bei den vorliegenden Daten von einem Kreis ausgegangen werden. Von einer Kugelhaftigkeit kann, sobald die anderen Organe die Form bedingen, nicht mehr oder nur sehr bedingt gesprochen werden. Für die Implementierung geeigneter mathematischer Verfahren und die Auswertung wird dann Mathematica 5.0 und SPSS 14 verwendet.



Abbildung 9.1: Bilderserie des Tumors nach der Magnetresonanztomographie

Im **1.Schritt** sind in allen Bildern einer Sequenz der jeweilige Tumor zu mar-

---

kieren und zu extrahieren. Dadurch ergibt sich allein schon bei der Erfassung durch die Unterschiede in der Qualität der Bilder eine je nach Bild verschiedene Unschärfe des Tumorrands.

Aus jedem der Bilder wird im **2.Schritt** der Tumor extrahiert. Es ergeben sich Schichten. Die Dicke der Schicht ergibt sich aus den Angaben der Magnetresonanztomographie<sup>1</sup>. Die Schichten werden zu einem dreidimensionalen Objekt zusammengefügt.

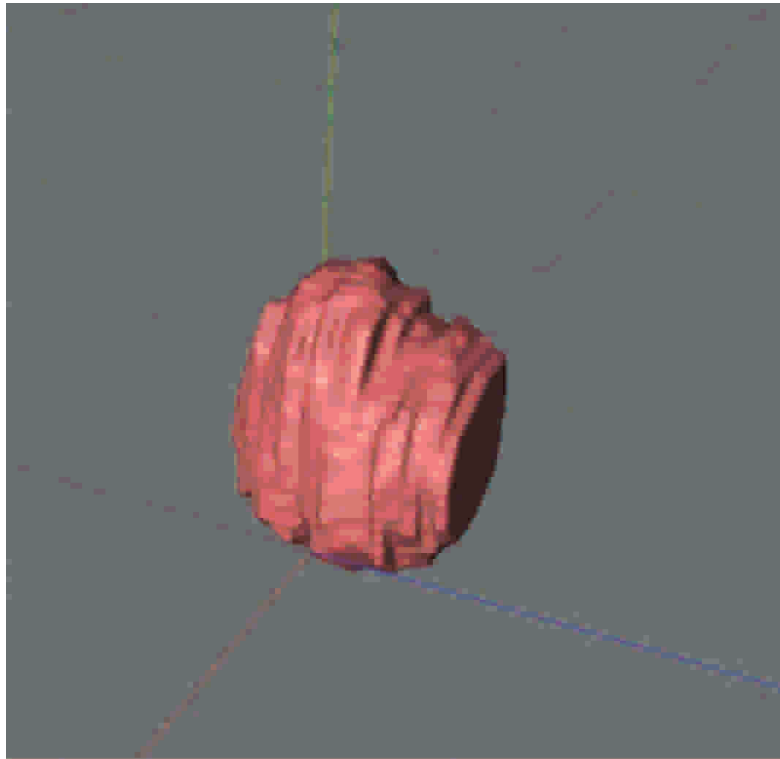


Abbildung 9.2: 3D-Darstellung des Tumors

Im **3.Schritt** werden alle Eckpunkte, die zur Beschreibungen des dreidimensionalen Objekts wie in Abbildung 9.10 benötigt werden, erfasst. Diese Eckpunkte werden allesamt auf den dreidimensionalen Schwerpunkt zentriert. Das Zentrum des nun zu untersuchenden Objekts liegt im Ursprung. Es gibt keine mehr zu beachtenden Unterschiede hinsichtlich der Lage des Tumors.

---

<sup>1</sup> Anmerkung: Vgl. Bildinformation in Voxel zu Pixel bei: [53]

---

Im **4.Schritt** wird der Tumor in den abgestumpften platonischen C60-Körper eingebettet, wie in Abbildung 9.3. Das Zentrum des platonischen Körpers stimmt mit dem Tumor überein.

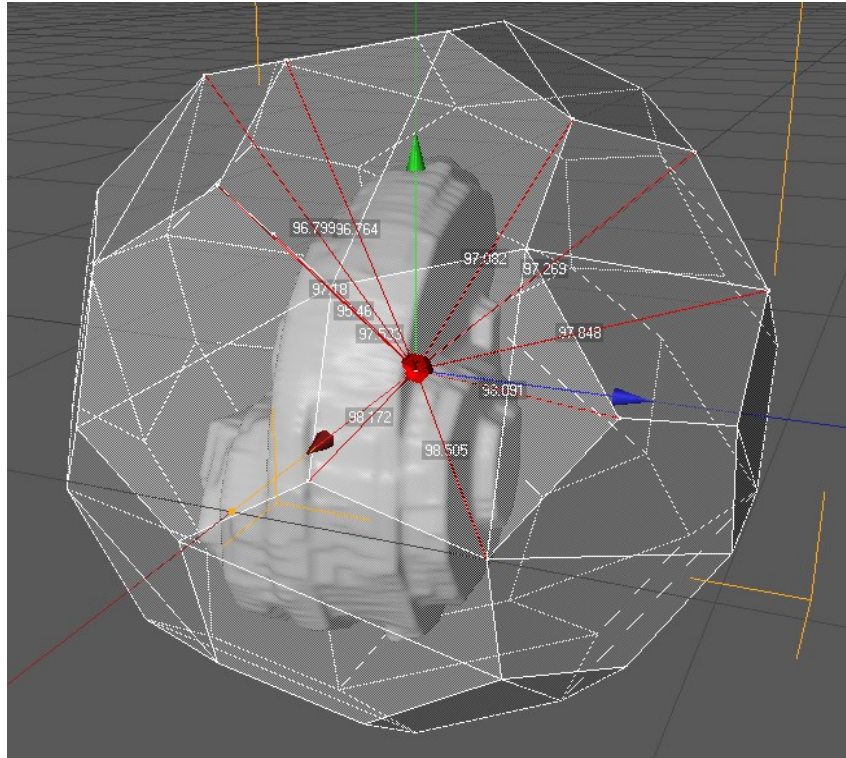


Abbildung 9.3: Einbettung des Tumors in den platonischen C60-Körper

Im **5.Schritt** werden die 60 Landmarken ermittelt. Die 60 Landmarken ergeben sich als Schnittpunkte der Linien vom Zentrum zum Eckpunkt mit der Oberfläche des Tumors. Dabei wird der nächstliegende Eckpunkt als Landmarke zur Beschreibung des Tumors gewählt.

In der vorher von mir im Rahmen der Diplomarbeit durchgeführten Untersuchung [32] wurde lediglich ein Bild aus der Sequenz gewählt, welches dem Schwerpunkt des Tumors am nächsten kam. Um den Schwerpunkt herum wurden 24 Landmarken erfasst, dargestellt in Abbildung 9.4.

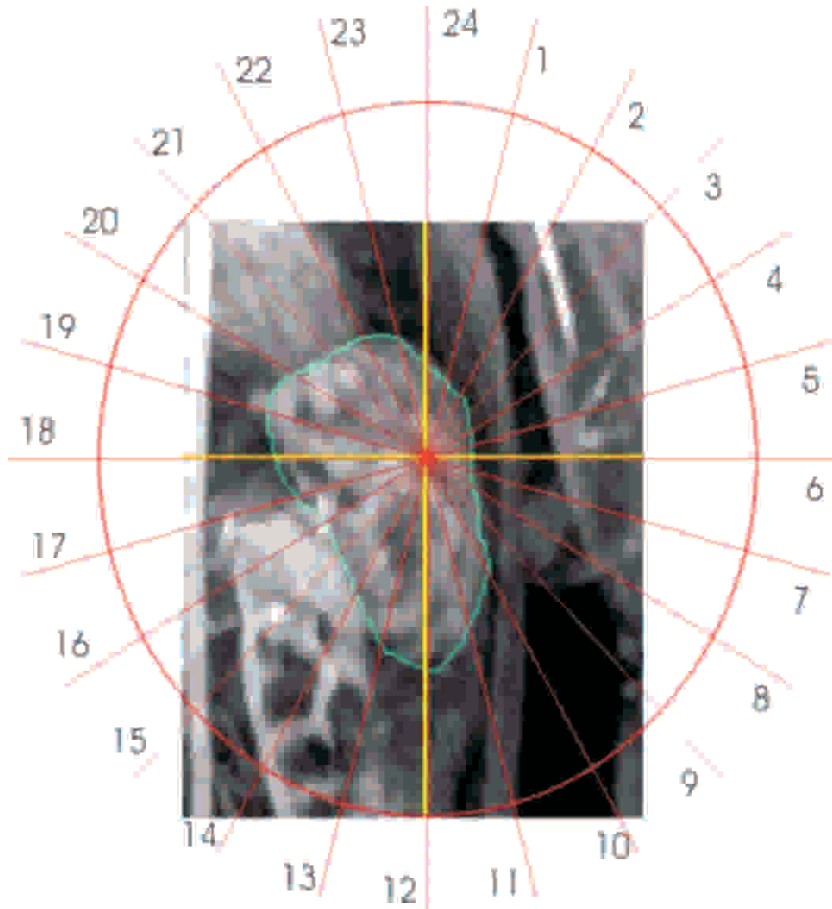


Abbildung 9.4: Einzeichnen des Randes und der Geraden in  $15^\circ$  Abständen durch den Ursprung

Die Reflexion wurde damals nur in der Transversalansicht des Tumors beachtet, nicht bereits wie in der vorliegenden Untersuchung bereits in der Frontalansicht. Grund dafür war die bestätigte Annahme, dass es sich bei den damals erhobenen Nephroblastomen im Zweidimensionalen noch um Kreise handelt [32].

### 9.1.2 Test auf Unterscheidbarkeit der Gruppen

Um die Unterscheidbarkeit der Gruppen sicher zu stellen, wird das Rauschen aus den Daten entfernt durch schrittweise Minimierung der Varianz in den Da-

ten. Dies wird für alle Patienten und erhobenen Landmarken in gleichem Maß durchgeführt. Nach Normierung und Zentrierung ist dies durch das Neuronale Netz entsprechend durchzuführen. Der Mittelwert wird aus den gewichteten Landmarken berechnet. Alle Objekte werden so transformiert, dass die Varianz mit jedem Schritt geringer wird.

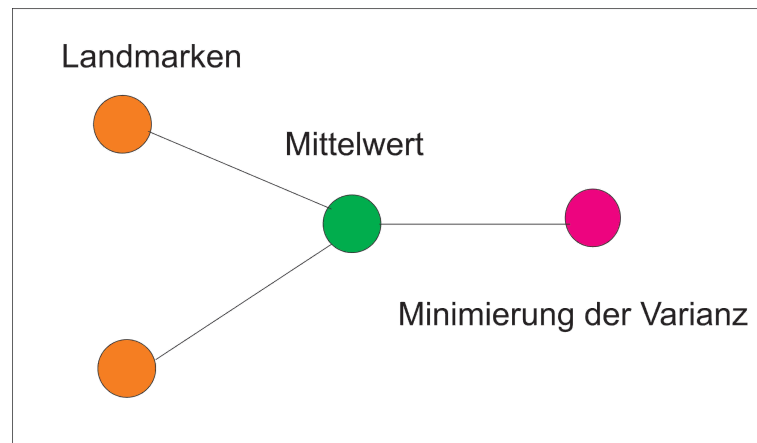


Abbildung 9.5: Minimierung der Varianz durch das Neuronale Netz

Für die Minimierung wird im Folgenden nach Normierung und Zentrierung nur eine Iteration durchgeführt. Der Mittelwert wird aus den Landmarken berechnet. Die Landmarken wiederum werden entsprechend für die Verringerung der Varianz um  $\frac{9}{10}$  verändert. Die Beweglichkeit der Landmarken nimmt dadurch ab. Die Schärfe des Bildes nimmt entsprechend zu.

Dabei ist zu beachten, dass die Minimierung nicht unbedingt zu einer Verbesserung der Differenzierbarkeit beiträgt. Bereits in den ersten Untersuchungen unserer Daten konnte für 74 vergleichbare Tumore (69 Nephroblastome und 5 Neuroblastome) die entsprechenden  $u_0$ -Werte für die Richtung Nephroblastome versus Neuroblastome berechnet werden.

Der  $u_0$ -Wert, der vom Neuronalen Netz berechnet wird, liegt zwischen 0 und 188. Es zeigt sich gerade beim Neuronalen Netz, dass nicht mit jedem Schritt der Minimierung der Varianz ein geringerer  $u_0$ -Wert erreicht wird. Der alternative Ansatz ist es zusätzlich zur Minimierung der Varianz bei Berechnung einen weiteren Parameter zu Differenzierung der Gruppen zu verwenden wie den  $u_0$ -Wert.

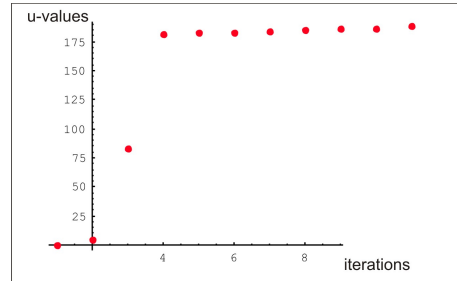


Abbildung 9.6: u-Werte nach jeder Iteration des Neuronales Netzes

### Mittlere Form in der Frontalansicht

Bei der Prüfung der mittleren Form der Nephroblastome versus der Nicht-Nephroblastome ergibt sich bereits für die mittlere Form der Nephroblastome in Abbildung 9.7 ein  $u_0$ -Wert von 270. Wird dieser verglichen mit weiteren 199 zufällig gezogenen 85 Objekten von insgesamt 108, so liegt der niedrigster  $u$ -Wert bei 275 und der höchste bei 721. Maximal möglich ist ein  $u$ -Wert von  $85 \cdot 23 = 1955$ . Der niedrige maximale  $u$ -Wert deutet auf eine geringe Variabilität der Tumore insgesamt hin. Der  $p$ -Wert liegt damit bei 0,005.

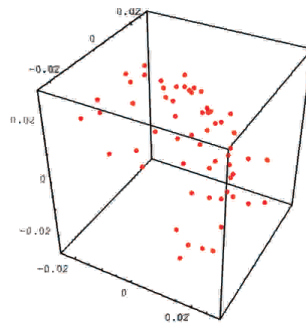


Abbildung 9.7: Mittlere Form der Nephroblastome

In der Richtung Nicht-Nephroblastome versus Nephroblastome ergibt sich ein  $u_0$ -Wert von 691 für die in Abbildung 9.8 abgebildete mittlere Form. Ebenfalls wird dieser mit 199 weiteren Werten verglichen und es ergibt sich der vierter Rang in eine Menge von 200  $u$ -Werten, die zwischen 427 und 1257 schwanken. Der  $p$ -Wert liegt damit bei  $p = 0,02$ . Es ergibt sich in beide Richtungen somit ein signifikantes ( $\alpha = 0,05$ ) Ergebnis.

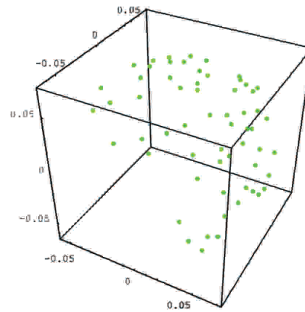


Abbildung 9.8: Mittlere Form der Nicht-Nephroblastome

Für die Gruppe der Nicht-Nephroblastome stellt sich die Frage, ob diese Gruppe noch in weitere Gruppen zu differenzieren ist (Klarzellensarkome, Rhabdoidtumore, Angiomyolipome, Neuroblastome und Nierenzellkarzinome). Die neun Neuroblastome erreichen einen  $u_0$ -Wert von 45. Der maximale  $u$ -Wert beträgt  $14 \cdot 9 = 126$ .

Ebenfalls wird dieser mit 199 weiteren Werten verglichen und es ergibt sich der erste bzw. der fünfte Rang in eine Menge von 200  $u$ -Werten, die zwischen 45 und 77 schwanken. Der  $p$ -Wert liegt damit bei  $p = [0,02; 0,025]$ .

Die übrigen Nicht-Neuroblastome haben einen  $u_0$ -Wert von 53. Mit 199 weiteren Werten ergibt sich der 75. bzw. der 136. Rang in eine Menge von 200  $u$ -Werten, die zwischen 49 und 80 schwanken. Der  $p$ -Wert liegt damit bei  $p = [0,375; 0,69]$ . Erwartungsgemäß taugt die "mittlere Form" der äußerst heterogenen Restgruppe nicht zur weiteren Differenzierung.

Für die Gruppe der fünf Nierenzellkarzinome ergibt sich ein  $u_0$ -Wert von 41. Bezogen auf die weiteren 199 Werte, die zwischen 36 und 50 liegen, theoretisch bei maximal 90, ergibt sich ein Rang von 84 bzw. 152. Der  $p$ -Wert liegt damit im Intervall  $[0,42; 0,76]$ . Die Gruppe der 18 Nicht-Nierenzellkarzinome innerhalb der Nicht-Nephroblastome hat einen  $u_0$ -Wert von 50. Die  $u$ -Werte schwanken zwischen 48 und 52 und der entsprechend der Ränge gebildete  $p$ -Wert liegt im Intervall  $[0,365; 0,78]$ .

Für die vier Klarzellensarkome ergibt sich ein  $u_0$ -Wert von 27. Die weiteren 199  $u$ -Werte bewegen sich zwischen 18 und 45, theoretisch liegt der maximale Wert bei 76. Es ergibt sich ein Rang zwischen 118 und 123 und ein  $p$ -Wert zwischen  $[0,59; 0,615]$ . Die Nicht-Klarzellensarkome erreichen einen  $u_0$ -Wert von 55. Ihr  $p$ -Wert liegt gemäß der Ränge zwischen  $[0,46; 0,655]$ .

Bei einer Fallzahl von weniger als oder gleich zwei Fällen stellt sich nicht nur die Frage der Generalisierbarkeit, sondern auch die Frage der Eignung des Algorithmus, der bei zwei Formen lediglich den bekannten arithmetischen Mittelwert berechnet. Daher werden die "mittlere Form" der Rhabdoidtumore, des Angiomyolipoms und eines nicht genauer diagnostizierten Nicht-Nephroblastom nicht



berechnet und geprüft.

Tabelle 9.1: Ergebnis der Unterscheidung der Nephroblastome von den Nicht-Nephroblastomen

Mengen		Unterscheidung der Arten			
Tumorart 1	Tumorart 2	$u_0$	$m_{=}$	$m_{<}$	$p - Intervall$
Nephroblastom	“Nicht-Nephroblastom”	270	0	0	0.005
“Nicht-Nephroblastom”	Nephroblastom	691	0	4	0.02

$m_{=}$ : Anzahl der Permutationen, bei denen der gleiche  $u$ -Wert sich ergibt;  $m_{<}$ : Anzahl der Permutationen, bei denen ein niedrigerer  $u$ -Wert vorliegt. Angabe des Intervalls für den  $p$ -Wert der Anfangshypothese über den niedrigsten und höchsten Rang des  $u_0$ - Wertes

Wird der Algorithmus mit dem Entscheidungsbaum und der Zielgrösse Diagnose kombiniert, so ist der erste Ast für die Unterscheidung Nephroblastome versus Nicht-Nephroblastome zu verwenden. Eine weitere Unterscheidung erfolgt dann in der Gruppe der Nicht-Nephroblastome und zwar nur hinsichtlich der Neuroblastome.

Demnach ist das relevanteste Merkmal die "mittlere Form" der Nephroblastome, gefolgt von der der Neuroblastome. Weitere mittlere Formen eignen sich nicht. Somit bestätigen sich die ersten Ansätze einer Focussierung auf die beiden Tumorarten Nephroblastome und Neuroblastome, wie dies bereits mehrfach von uns vermutet wurde [33].

Der Test der "mittleren Form" der Nephroblastome wird zusätzlich unter Beachtung der Varianz der Distanzen durchgeführt. Dabei wird die Bedingung für eine Erhöhung des  $u$ -Wertes verändert. Das Objekt, welches zur Gruppe der Nephroblastome gehört, muss eine höhere Distanz von der "mittleren Form" und die addierte 20%ige Standardabweichung aufweisen. Es ergibt sich ein  $u_0$ -Wert der Nephroblastome von 217 und entsprechend wieder ein  $p$ -Wert von 0,005.

## Varianz

Wie bereits im Ansatz des einstigen Tests von Ziezold(1994) [90] beschrieben, kann sich das Ergebnis auch als Folge der unterschiedlichen Varianz in den beiden Gruppen ergeben. Dies ist allein durch die erstmalige Einbeziehung der Standardabweichung nicht mehr der Fall.

Genauso kann auch die Varianz an sich geprüft werden. Dafür wird die mehrdimensionale Varianz verwendet. Sollte sich kein Unterschied in der Varianz

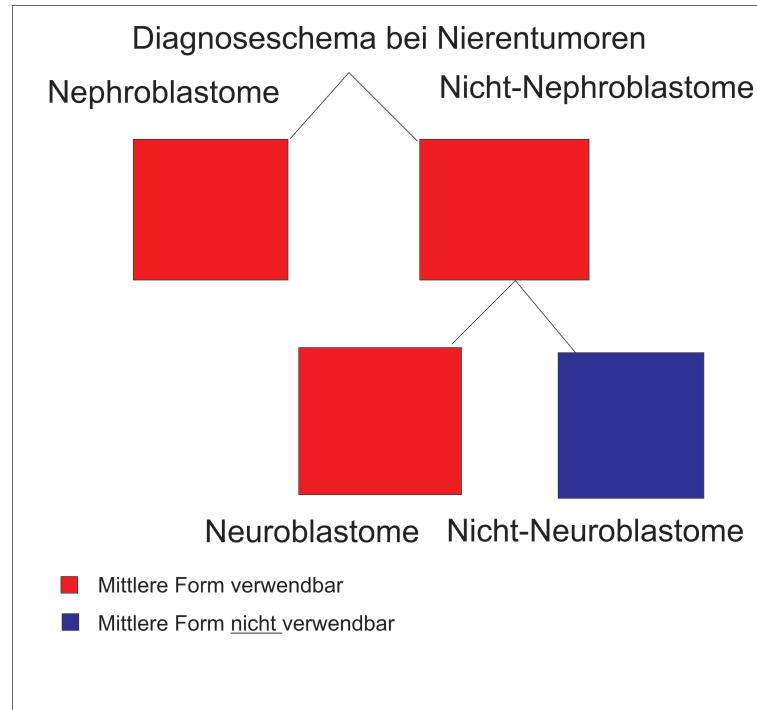
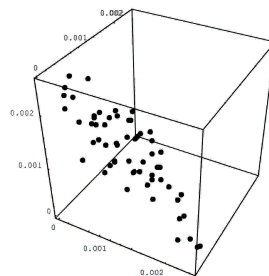


Abbildung 9.9: Diagnoseschemata bei Nierentumoren von Kleinkindern

ergeben, spricht dies dafür, dass der Unterschied der mittleren Form ein Ergebnis der Differenzierbarkeit beider Gruppen darstellt.

Die Varianz ist bei einem  $F$ -Wert von 1,1157 nahezu gleich in beiden Gruppen. Der Rang des  $F$ -Wertes liegt außerhalb der Möglichkeit die Nullhypothese einer Gleichheit der Varianzen zu verwerfen.

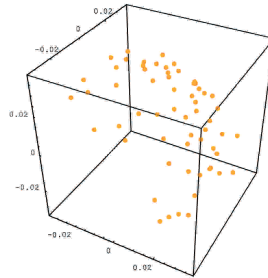
**Beispiel 36.** Zur Veranschaulichung der Varianz und der mittleren Form werden alle Tumordaten verwendet



Mehrdimensionale Varianz bei den erhobenen Tumoren

---

Im Vergleich dazu die "mittlere Form" aller Tumore:



Mittlere Form bei den erhobenen Tumoren

### 9.1.3 Test der Landmarken zur Unterscheidung der Gruppen

#### Wilcoxon-Test

Beim Wilcoxon-Test ergibt sich ein  $u_0^+$ -Wert von 0 für die "mittlere Form" der Nephroblastome. Verglichen mit 199 weiteren Werten ergibt sich ein  $p$ -Wert von  $\frac{1}{200}$ . Maximal kann der Wert 3600 erreicht werden. Das wäre der Fall, wenn alle Abweichungen der Landmarken der zur Bestimmung der "mittleren Form" verwendeten Objekte höhere Werte erreichen als die Mittelwerte der Distanz derer, die nicht zur "mittleren Form" verwendet wurden. Der Wert liegt in der Zufallsstichprobe zwischen 0 und 3600. Es handelt sich um ein signifikantes Ergebnis der "mittleren Form" der Nephroblastom nach dem so verwendeten Wilcoxon-Test.

#### Explorativer Permutationstest

Um aus der Menge von 60 Landmarken relevante Landmarken zu erkennen, werden zufällig 5 Landmarken aus der Menge der 60 Landmarken gezogen und die Fünferkonfiguration mit dem niedrigsten  $u_0$ -Wert gewählt. Dann wird die Fünferkonfiguration analog zum Test der "mittleren Form" für die Differenzierung der Objektgruppen geprüft. Wenn von einer Gleichverteilung der  $u$ -Werte ausgegangen wird, kann der entsprechende  $p$ -Wert aus dem Rang in der Liste der  $u$ -Werte berechnet werden.

Bei 200 zufällig gezogenen Fünferkonfiguration, ergibt sich mit einem  $u$ -Wert von 108 die Konfiguration {25, 40, 41, 43, 44}. Damit liegen die ausgewählten relevanten Landmarken im oberen Bereich des Körpers. Die Landmarken Richtung Niere sind bei dieser Auswahl als weniger relevant einzuschätzen. Das Ergebnis in diesem Fall ist in Folge der Beachtung der Reflexion direkt in die Anwendung zu übertragen.

### 9.1.4 Ergebnisse der Klassifikation

Bei der Klassifikation wird aufgrund der Gruppengröße nur zwischen Nephroblastomen und Nicht-Nephroblastomen unterschieden. Erst mit der Klassifikation anhand der vorliegenden Daten gelangt der Anwender auf die Ebene einer Diagnose.

#### Entscheidungsbäumen

Die Entscheidungsbäume sind mit dem "Exhaustive-Chaid"-Algorithmus, implementiert in SPSS 14, in der Lage, jeweils schrittweise bivariat die Kategorien der erklärenden Variablen so zusammenzufassen, dass eine Klassifikation anhand der Variablen "Alter", "Lage des Tumors" (Links/Rechts) und der Distanz von der mittleren Form der Nephroblastome zu der Diagnose "Nephroblastom" bzw. "kein Nephroblastom" vorgenommen werden kann.

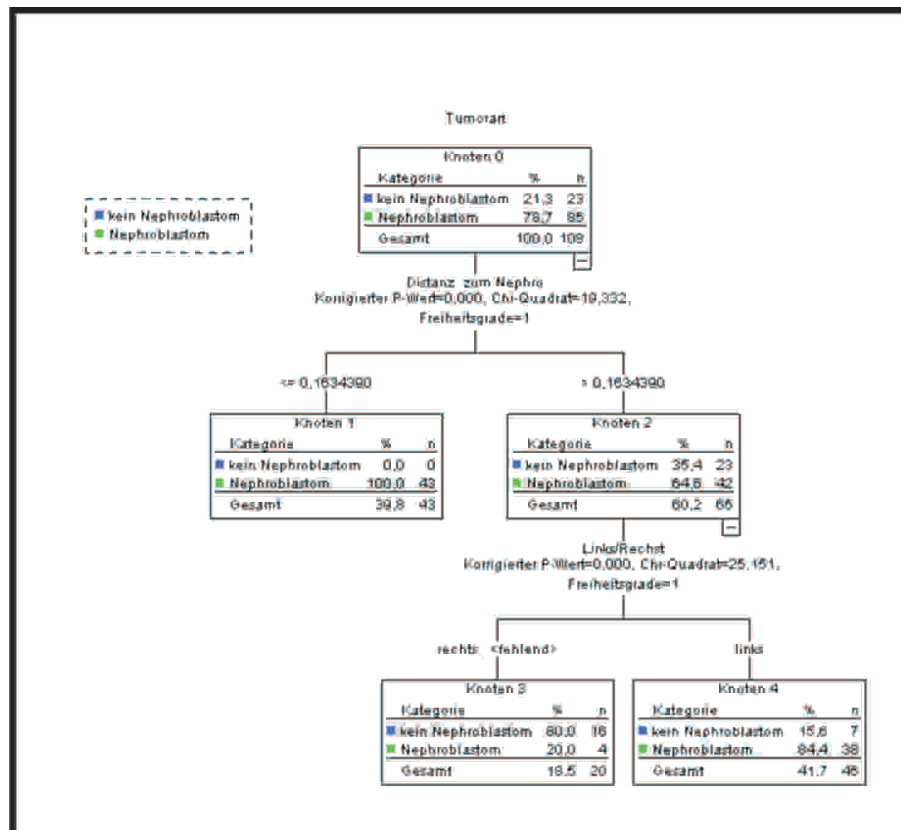


Abbildung 9.10: Klassifikation mittels der Entscheidungsbäume aller Tumore anhand der Variablen "Alter", "Lage des Tumors" und "Distanz von der mittleren Form"

---

Die Distanz von der mittleren Form ist nach den Entscheidungsbäume die entscheidende erklärende Variable für die Zuordnung zu der Diagnose. Die Auswertung der Lage des Tumors weist erheblich mehr Nephroblastome auf der linken Seite als auf der rechten hin.

Insgesamt werden 16 von 23 Nicht-Nephroblastome korrekt zugeordnet und 81 von 85 Nephroblastome. 89,1% können somit korrekt zugeordnet werden. Das Verfahren ist darüber hinaus in der Lage, auch mit fehlenden Werten durch die Einordnung in die Kategorien sinnvoll und für den Anwender nachvollziehbar umzugehen.

### Diskriminanzanalyse

Die Diskriminanzanalyse nach Fisher ordnet anhand der Distanz in der Frontalansicht zur "mittleren Form" der Nephroblastome, des Alters und der Lage (Links-/Rechts) 19 von 21 der Nicht-Nephroblastome korrekt zu, sowie 82 von 82 Nephroblastomen. Insgesamt werden 98,1% der Patienten richtig zugeordnet.

Die standardisierten Funktionskoeffizienten betragen für das Alter 0,109, für die Lage 1,693 und für die Distanz 1,906. Da die Gruppenmittelwert für Nephroblastome bei 3,255 und für Nicht-Nephroblastome bei  $-0,834$  liegen, bedeutet dies, dass mit zunehmenden Alter, mit der rechten Lage des Tumors und einer stärkeren Entfernung von der "mittleren Form" der Nephroblastome eher von einem Nicht-Nephroblastom auszugehen ist. Das Ergebnis des Modells ist signifikant (Wilks  $\lambda = 0,216$ ,  $df = 3$ ,  $\chi^2 = 131,982$ )<sup>2</sup>.

### Logistische Regression

Die Logistische Regression hat analog zur Diskriminanzanalyse die Distanz zur "mittleren Form" der Nephroblastome, das Alter und die Lage des Tumors zur Klassifikation zur Verfügung. 20 von 21 der Nicht-Nephroblastome werden korrekt zugeordnet, sowie 81 von 82 Nephroblastomen. Insgesamt werden ebenfalls 98,1% der Patienten richtig zugeordnet.

Analog zur Diskriminanzanalyse stellt sich die Veränderung mit der Distanz zur "mittleren Form" der Nephroblastome (unstandardisiertes Chancenverhältnis: 0,00075) ein, mit der Lage (unstandardisiertes Chancenverhältnis: 0,00017 und schließlich mit dem Alter in Tagen (unstandardisiertes Chancenverhältnis: 0,998834). Signifikant ( $p < \alpha = 0,05$ ) sind im Modell die Lage des Tumors und die Distanz. Das Modell erklärt mit Nagelkerkes  $R^2 = 0,958$  die Unterschiede zwischen den beiden Gruppen<sup>3</sup>.

---

<sup>2</sup> [3], S.173

<sup>3</sup> [3], S.133

---

## Mehrschichtiges Perzeptronennetz in Kombination mit der Formanalyse

Zunächst wird analog zu den obigen Verfahren das Neuronale Netz zur Klassifikation verwendet. Es ergibt sich nach wenigen Durchgängen nur noch eine Fehlklassifikation (Anzahl der versteckten Neuronen=1 und Iterationen =10). Der 89.te Patient wird fälschlich der Kategorie "Nicht-Nephroblastom" zugeordnet.

Bei der Anwendung Neuronaler Netze sind die Einstellungsparameter zu beachten. Bei bestimmten Einstellungen bleibt das Neuronale Netz in einem lokalen Minimum hängen. Die Gewichtsveränderungen reichen dann nicht mehr aus, um das lokale Minimum zu verlassen. Das Potential Neuronaler Netze, eine Zuordnung zwischen Ein- und Ausgabe zu finden, liegt generell bei entsprechender Betrachtung nichtlinearer Abhängigkeiten und Wechselwirkungen über dem der Diskriminanzanalyse und der Logistischen Regression.

Das bereits in Abschnitt 6.4 beschriebene und nicht im Vergleich zu den obigen Verfahren stehende Vorgehen wird nun nur für die Distanz zur "mittleren Form" der Nephroblastome angewandt. Angesichts der Fallzahl der übrigen Diagnosen ist auch bei der Verwendung Neuronaler Netze eine Aussage, die bei den einzelnen Arten der Nicht-Nephroblastome über die vorliegenden Daten hinaus zu verwenden ist, nicht zu erwarten.

Dabei werden alle 60 Landmarken einzeln gewichtet und der Anwendung genüge getan, die davon auszugehen hat, dass nicht jede der Landmarken eine gleiche Relevanz für die Klassifikation aufweist. Die Gewichtung der Landmarken kann für den Anwender zur Deutung herangezogen werden. Um Veränderungen der Gewichtungen je nach zufällig initialisierten Startwerte zu vermeiden, wird aufgrund fehlender Kenntnis des Tumorgewebes von einer gleichen Gewichtung ausgegangen. Bei diesem Vorgehen können nach 20 Iterationen 98 Patienten korrekt zugeordnet werden. Alle Landmarken in Form des summierten Abstandes von der "mittleren Form" werden in der endgültigen Zuordnung in gleichem Umfang gewichtet.

Neuronale Netze können im Gegensatz zu den obigen Verfahren auch mit einer grösseren Datenmenge an erklärenden Variablen umgehen, da sie keinen Verteilungsannahmen zu genügen haben. Sie sind jedoch auch nicht in der Lage, die Güte des Modells anzugeben. Diese ist nur durch entsprechende Datensplits möglich. Angesichts der Größe des Datensatzes und der Heterogenität der Nicht-Nephroblastome wird kein Datensplit vorgenommen.

## P-KFA

Die P-KFA ist bei der Klassifikation insbesondere dann von Nutzen, wenn zusätzlich zu den stetigen auch diskrete Merkmale zu verwenden sind. Die Logistische Regression, die Diskriminanzanalyse und das Mehrschichtige Perzeptro-

nennetz verarbeiten die nominal skalierten Variablen wie die stetigen. In dem Fall der P-KFA wird das Alter in Kategorien verwendet. Während laut der Medizin bei der Diagnose der Nierentumore im Kindesalter das Geschlecht unbeachtet bleiben kann [9], ist gerade bei der Diagnose der Nierenzellkarziome das Alter zu berücksichtigen. Um nicht zu der Entscheidung zu gelangen, alle Patienten ab einem bestimmten Alter zu einer der Kategorien von Tumorarten zuzuordnen, wird die Kombination des Alters mit der Distanz verwendet. Darüber hinaus wird die Lage des Tumors (Links : "1", Rechts: "2") verwendet, da nicht alle Organe wie die Milz symmetrisch im Körper auf beiden Seiten anzutreffen sind.

Da bei Kindern zwischen dem 9. und 15. Lebensjahr von Nierenzellkarzinomen ausgegangen werden kann [78], wird die Schwelle bei 9 Jahren angesetzt und die Schwelle für die Distanz zur mittleren Form bei 0,5. Das bedeutet, dass alle Werte über bzw. gleich der Schwelle der Kategorie "2" zugeordnet werden und alle Werte unterhalb der Kategorie "1". Es ergibt sich folgende Tabelle unter Verwendung des Programms [75]:

Tabelle 9.2: Patientenprofile

Patientenprofil		Bewertung
	Häufigkeit	$p$ -Wert
111x1	0	1
111x2	2	0,039977
112x1	0	1
112x2	2	0,039977
121x1	2	0,632210
121x2	0	1
122x1	1	0,992480
122x2	2	0,104891
211x1	7	0,985249
211x2	5	0,065343
212x1	0	1
212x2	2	0,039977
221x1	29	0,000371
221x2	0	1
222x1	43	0,175747
222x2	8	0,122420

*Profil: alter, distanz, lage x nephroblastom*

---

Nach der Bonferoni-Adjustierung  $\alpha/(2^4) = 0,003125$  ergibt sich nur das Profil 221 als signifikant der Kategorie "1" zuzuordnen. Patienten mit einem Alter über 9 Jahre, einer Distanz von der mittleren Form geringer als 0,5 und auf der linken Seite gelegen, sind eher verglichen mit der Häufigkeit des Auftretens der Tumorarten als Nephroblastome zu klassifizieren. Damit bietet dieses Verfahren erstmalig eine Möglichkeit die Gruppe der älteren Patienten zu differenzieren.

## 9.2 Räumliche Analyse bei Serienmördern

### 9.2.1 Erfassung der Daten

Der Raum, der durch das Handeln des Täters geprägt wird, ist durch drei zeitlich aufeinanderfolgende Handlungen innerhalb des jeweiligen Tötungsdeliktes beschrieben:

1. Begegnungsort mit dem Opfer
2. Tatort
3. Leichenfundort

Der Ausgangspunkt für die Orte ist der Ankerpunkt des Täters. Der Ankerpunkt ist durch den Lebensraum des Täters bestimmt. Von diesem Punkt aus kann er seine Tat begehen. Ankerpunkte können die Wohnung des Täters, die Arbeitsstelle, sowie die Wohnung seiner Primärfamilie (Kinder, Eltern, Partnerin) sein. Es kann sich aber auch um Ankerpunkte wie ein Ferienhaus oder ihm aus Kindheit/Jugend usw. noch bekannte Orte handeln.

Mit den bisher verwendeten mathematischen Verfahren wie dem  $\chi^2$ -Test wird lediglich der Zusammenhang zwischen dem Abstand eines Ortes vom Ankerpunkt in Kategorien und den Ausprägungen des Täters untersucht. Es werden nicht alle drei Orte gleichzeitig zur Ermittlung eines spezifischen Vorgehens des Täters verwendet und die Bildung von diskreten Kategorien führt zu einem Informationsverlust der stetig erfassten Merkmale.

Unbeachtet bleibt in der Regel bei den Untersuchungen von Serienmördern, dass das Potential des Täters, sich im Raum zu bewegen, mit den anderen Tätern für eine Untersuchung vergleichbar zu sein hat. Werden die Daten unverarbeitet übernommen, bleiben auch die Unterschiede in der Mobilität des Täters erhalten. Das bedeutet, dass bei der Gewinnung einer Aussage über eine bestimmte Tätergruppe diese Aussage letztendlich nicht unbedingt auf eine die Tätergruppe bestimmende Ausprägung zurückzuführen ist, sondern Folge der Möglichkeit und der Fähigkeit, sich im Raum zu bewegen.

Die Daten der 49 Serienmörder liegen bereits zentriert auf den Ankerpunkt vor. Um die eindimensionalen Entfernungen vom Ankerpunkt zweidimensional darzustellen, wird die  $y$ -Koordinate des Begegnungsortes auf  $-1$  gesetzt, die des Tatortes auf  $0$  und die der  $x$ -Achse des Leichenfundorts auf  $0$ . Dadurch ergibt sich selbst im Falle eines räumlichen Zusammenfallens aller drei Handlungsorte ein räumliches Profil, ohne eine Veränderung der numerischen Werte wie in



---

der ersten von mir vorgenommenen Untersuchung [31]. Im Falle eines gleichen Abstands von Ankerpunkt zum Begegnungsort und zum Leichenfundort, sowie dem Zusammenfallen von Ankerpunkt und Tatort liegen alle drei Punkte auf eine Linie.

## 9.2.2 Auswertung

### 9.2.3 Test auf Unterscheidbarkeit der Gruppen

Die Gruppen von Serienmördern ergeben sich aus den Eigenschaften des Täters, der Tat und des Tathergangs. Die Polizei möchte anhand des räumlichen Vorgehens bei der ersten Tat Rückschlüsse ziehen können. Zumindest haben alle sexuell motivierten Serienmörder zwei und mehr voneinander unabhängige Taten begangen. Demnach wäre es hilfreich schon durch das Begehen der ersten Tat Eigenschaften des Täters, seiner Tat und seines Modus operandi zu ermitteln.

Zu erwähnen ist bei der vorgenommenen Untersuchung, dass der Polizei nicht unbedingt in jedem Fall bei der Ermittlung bereits alle drei Orte bekannt sind.

#### Tat 1

Zunächst ist die Eigenschaft der Beziehung zu sehen. Für die Tat 1 ergibt sich bei einer Beziehung zum Opfer ( $n = 15$ ) ein  $u_0$ -Wert von 277. Die weiteren 199  $u$ -Werte schwanken zwischen 282 und 374. Der Rang des  $u_0$ -Wertes ergibt sich demnach zwischen 1 und 2 und der  $p$ -Wert  $[0,005; 0,01]$ .

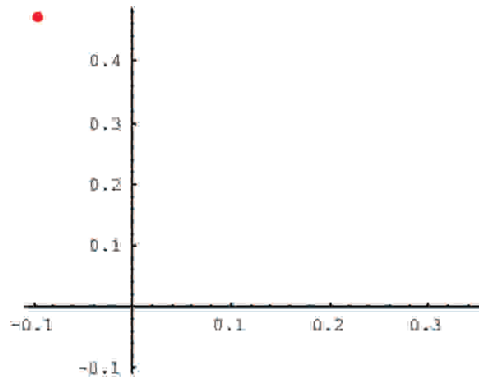


Abbildung 9.11: Mittleres räumliches Profil bei der Tat Nr. 1 in der Gruppe der Serienmördern mit Beziehung zum Opfer

Für diejenigen, die eine Beziehung zum Opfer hatten ( $n = 34$ ) ergibt sich ein  $u_0$ -Wert von 147. Die weiteren 199  $u$ -Werte schwanken zwischen 146 und 188.

---

Der Rang des  $u_0$ -Wertes ist 2 und der  $p$ -Wert 0,01. Im Bild ist zu erkennen, dass Tatort und Begegnungsort innerhalb dieser Untersuchung nahe beieinander liegen, der Leichenfundort sich weiter davon entfernt.

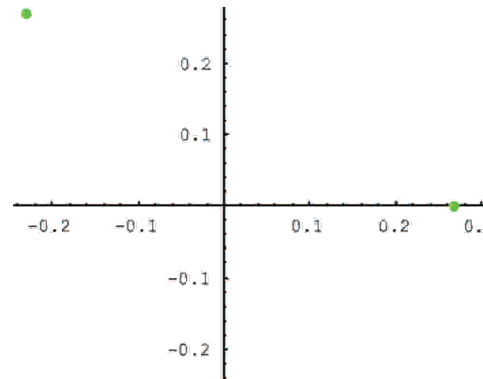


Abbildung 9.12: Mittleres räumliches Profil bei der Tat Nr. 1 in der Gruppe der Serienmördern ohne Beziehung zum Opfer

---

Die 6 jugendlichen Straftäter (jünger als 18 Jahre) haben einen  $u_0$ -Wert von 108. Die weiteren 199  $u$ -Werte schwanken zwischen 105 und 116. Der Rang des  $u_0$ -Wertes ergibt sich zwischen 12 und 15 und der  $p$ -Wert  $[0,06; 0,075]$ . Für die übrigen 43 Straftäter (älter als 18 Jahre) ergibt sich ein  $u_0$ -Wert von 85. Die weiteren 199  $u$ -Werte schwanken zwischen 95 und 165. Der  $p$ -Wert ergibt sich zu 0,005.

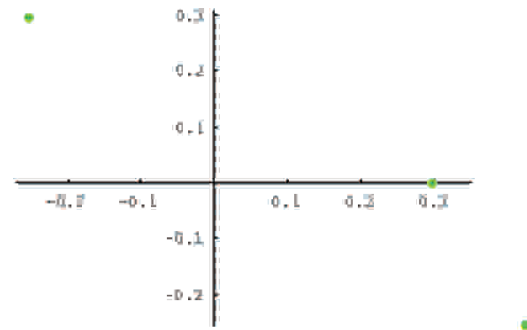


Abbildung 9.13: Mittleres räumliches Profil bei der Tat Nr. 1 in der Gruppe der Serienmörder älter als 18 Jahre

Für die übrigen 43 Straftäter (älter als 18 Jahre) ergibt sich ein  $u_0$ -Wert von 85. Die weiteren 199  $u$ -Werte schwanken zwischen 95 und 165. Der  $p$ -Wert ergibt sich zu 0,005.

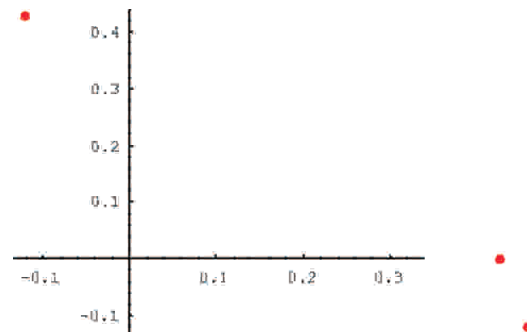


Abbildung 9.14: Mittleres räumliches Profil bei der Tat Nr. 1 in der Gruppe der Serienmörder jünger als 18 Jahre

Das bedeutet, dass hinsichtlich des räumlichen Profils von jüngeren Tätern von einer tendenziellen Differenzierbarkeit ( $\alpha = 0,1$ ) anhand des so im Zweidimensionalen abgebildeten räumlichen Profils gesprochen werden kann, während

bei sexuell motivierten Serienmördern älter als 18 von einem signifikanten Profil gesprochen werden kann.

## 9.3 Sensordaten elektronischer Nasen

### 9.3.1 Messungen

Es stehen für die Untersuchung der Geruchsquantität 115 olfaktometrische Messungen und die entsprechend zugeordneten Sensorsignale an einem Standort zur Verfügung. Für die Geruchsqualität sind 86 Messungen an unterschiedlichen Standorten vorhanden.

### 9.3.2 Test auf Unterscheidbarkeit der Gruppen

#### Mittlere Form

Hinsichtlich der *Geruchsquantität* wird die Schwelle  $1000 \text{ GE}/m^3$  verwendet und die Profile für unter  $1000 \text{ GE}/m^3$  und "über/gleich  $1000 \text{ GE}/m^3$ " hinsichtlich ihrer Differenzierbarkeit untersucht [?]. Für die Geruchsqualität werden 34 Messungen an unterschiedlichen Standorten hinsichtlich ihres Geruchsprofils differenziert. Es ergibt sich für die Geruchsquantität "unter  $1000 \text{ GE}/m^3$ " ein  $u_0$ -Wert von 1382 ( $n = 73$ ). Die weiteren 199  $u$ -Werte schwanken zwischen 1449 und 1618. Der Rang des  $u_0$ -Wertes ist 1 und der  $p$ -Wert 0,005.

Es ergibt sich für die Geruchsquantität "über/gleich  $1000 \text{ GE}/m^3$ " ein  $u_0$ -Wert von 1282 ( $n = 42$ ). Die weiteren 199  $u$ -Werte schwanken zwischen 1405 und 1636. Der Rang des  $u_0$ -Wertes ist 1 und der  $p$ -Wert 0,005. Hinsichtlich der Geruchsquantität wird die Schwelle  $1000 \text{ GE}/m^3$  verwendet und die Profile für unter  $1000 \text{ GE}/m^3$  und "über/gleich  $1000 \text{ GE}/m^3$ " hinsichtlich ihrer Differenzierbarkeit untersucht [?].

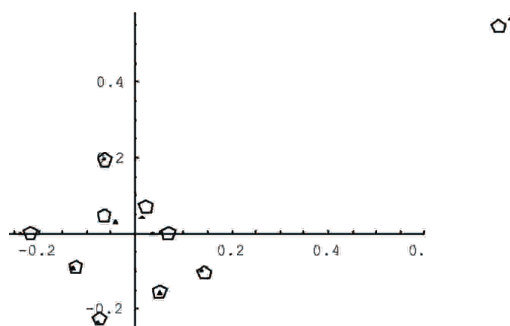


Abbildung 9.15: Mittleres Geruchsprofil bei über  $1000 \text{ GE}/m^3$  (Dreieck) und alle unterhalb

Anhand des Bildes 9.15 wird erkennbar, dass nur für bestimmte Sensoren eine Verwendbarkeit zur Bestimmung der vorgenommenen Kategorien von Geruch gilt. Die Profile sind sich aufgrund ihrer Messung an dem selben Standort sehr ähnlich. Trotzdem ermöglichen sie auch nach Normierung eine Aussage zur Geruchsquantität. Dies lässt darauf schliessen, dass nicht nur der Ausschlag von Sensorsignalen Aussagen zur Geruchsbelastung rechtfertigt, sondern bereits die Konfiguration der Sensoren.

Für die *Geruchsqualität* werden 86 Messungen an unterschiedlichen Standorten hinsichtlich ihres Geruchsprofils differenziert. Bezüglich der Geruchsqualität sind die Standorte "Bottrop 1" ( $n = 24$ ) und "Hilter" ( $n = 19$ ) von den übrigen Geruchsprofilen zu differenzieren. Für den Standort "Bottrop 1" ergibt sich ein  $u_0$ -Wert von 190. Die weiteren 199  $u$ -Werte schwanken zwischen 297 und 895. Der Rang des  $u_0$ -Wertes ist 1 und der  $p$ -Wert 0,005. Für das mittlere "Geruchsprofil" der Standorte "Nicht-Bottrop 1" ergibt sich ein  $u_0$ -Wert von 632. Die weiteren 199 Werte liegen zwischen 671 und 1051. Es ergibt sich in dieser Richtung ebenfalls ein Rang von 1 und ein  $p$ -Wert von 0,005.

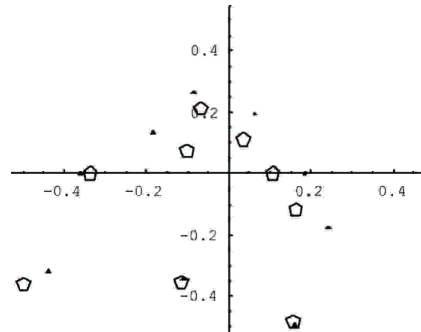


Abbildung 9.16: Mittleres Geruchsprofil bei Bottrop 1 (Dreieck) und allen anderen Standorten

Es zeigen sich deutlich zwei verschiedene Geruchsprofile. Für "Hilter" ergibt sich ein  $u_0$ -Wert von 427. Die weiteren 199  $u$ -Werte schwanken zwischen 421 und 491. Der Rang des  $u_0$ -Wertes ist 11 oder 12 und der  $p$ -Wert liegt zwischen  $[0,055, 0,06]$ . Damit ist das Geruchsprofil von "Hilter" nicht signifikant unterscheidbar von den übrigen Geruchsprofilen.

---

# Kapitel 10

## Fazit

### 10.1 Formanalyse

In der vorliegenden Arbeit wurde die Formanalyse auf drei unterschiedliche Disziplinen angewandt. In allen drei Disziplinen ergaben sich nach Entwicklung geeigneter mathematischer Verfahren für den Anwender verwendbare Ergebnisse. Das Vorgehen bei der Datenerfassung und die Datenverarbeitung in den jeweiligen Disziplinen bieten Ansätze für weitere Anwendungen.

Bei den Nierentumoren zeigte sich darüber hinaus die Notwendigkeit mittels Neuronaler Netze die Beweglichkeit der Landmarken durch die Minimierung der Varianz zu verringern. Ohne diese Vorverarbeitung sind keine verwendbaren Ergebnisse im Bereich der Nierentumore allein aufgrund der Unschärfe des Bildmaterials zu erzielen.

Insgesamt wurde ein gänzlich neues Gebiet in dieser Arbeit mit der Anwendung auf dreidimensionale Daten angegangen. Die beschriebene dynamische Formanalyse bietet im Weiteren den Ausblick für die vierdimensionalen Daten. Bei diesem Ansatz sind die Art der Abhängigkeiten zwischen Zeit und Landmarken noch näher zu untersuchen und entsprechend anzugehen.

Generell sind in der angewandten Mathematik das Problem der Datenerfassung, des Messfehlers und die Wahl eines geeigneten Modells zu sehen. Nur wenn alle Bereiche angegangen werden, sind entsprechend verwendbare Ergebnisse in der Disziplin zu erwarten.

### 10.2 Datenerfassung

Vor der Auswertung der Daten in den jeweiligen Disziplinen hat sich der Anwender über die Art der Erfassung und die Vergleichbarkeit Gedanken zu machen. Gerade die Formanalyse zeigt, wie wichtig der richtige Ansatz der Datenerfassung ist, um entsprechend verwertbare Ergebnisse generieren zu können. Das Vorgehen bei der Erfassung der Daten ist einer der wesentlichen Punkte der vorliegenden Arbeit.

---

Meist gibt es zu jedem Anwenderproblem geeignete mathematische Verfahren oder zumindest Ansätze für ein solches. Nur fehlt es an Daten, die für eine Auswertung geeignet sind.

## 10.3 Ungewissheit der Messung

Der Ungewissheit der Messung kann gerade durch die Normierung und Zentrierung entgegengewirkt werden. Wenn der Fehler sich auf alle Landmarken gleichermaßen auswirkt, hat er keinerlei Einflusss auf die Deutung der Form. Darüber hinaus ist der Test durch die Einbeziehung der Varianz entsprechend erweitert worden. Selbst unter Beachtung der Varianz ergeben sich brauchbare Ergebnisse.

Bei der Umrandung der Tumore kann genauso wenig von einer Exaktheit gesprochen werden wie bei dem rekonstruierten Tathergang eines Serienmörders oder bei der Messung einer elektronischen Nase.

Das Gewebe des Tumors ist in einigen Bildern sehr unscharf und hängt vom Kontrastmittel und vom Vorgehen bei der Bildgebung ab, die Rekonstruktion des Tathergangs von den Beweisen und der Bereitschaft des Täters, Auskunft über seine Tat zu geben und die Messung der elektronischen Nasen von den Störstoffen in der Luft und den technischen Komponenten.

Allen aufgrund der Unsicherheit in der Messung kann es zu Ergebnissen kommen, die vom Anwender nur bedingt erklärt werden können wie im Beispiel der von mir vorgenommenen Auswertung der Rückfallverläufe ohne Beachtung oder Herausrechnen der Haftzeiten, die nur unzureichend dokumentiert werden [41].

## 10.4 Mathematische Verfahren

### 10.4.1 Selektion der Modelle

Die angewandte Mathematik hat die Aufgabe für die jeweiligen Probleme hinsichtlich der Daten und der Fragestellung geeignete Modelle zu entwickeln. Die Selektion eines Modells hängt von den Annahmen ab. Je weniger der Anwender weiß, desto eher hat er die Annahmen des jeweiligen Modells zu prüfen.

Insoweit der Anwender nur ein Modell anhand aller Daten entwickelt, ist er im Bereich der "Erklärung". Vorhersage ist nur dann möglich, wenn die Annahmen des Modells zutreffen. Letztendlich ist dies anhand unbekannter Daten zu prüfen.

Die verwendeten Tests nehmen die in den Daten vorliegenden Verteilungen. Sie sind besonders für den Fall geeignet, dass die Anzahl an Patienten nicht ausreicht, um Verteilungen zu prüfen. Nachteilig ist die erhebliche Rechendauer der verwendeten Permutationstests.

Die Formanalyse hat sich als in vielen Bereichen anwendbar erwiesen. Sowohl im Bereich der Medizin, der Kriminologie und der Auswertung der Sensordaten konnte die vorliegende Arbeit die Verwendbarkeit der Formanalyse anhand der



---

Beispiele bestätigen und die Ansätze entsprechend der Disziplinen weiterentwickeln, sowie auf weiteres Potential wie in der dynamischen Formanalyse bei der Auswertung hinweisen.

Im Allgemeinen ist immer zu beachten bei der Auswahl des Verfahrens die Eignung hinsichtlich der einzugehenden Voraussetzungen (Verteilungsannahmen, Skalierung etc.) und hinsichtlich der gewünschten und benötigten Ergebnisse (Bestimmung einer Zielgrösse oder Gruppenbildung aus den Daten heraus), sowie die Anpassung des Verfahrens auf das Problem (Beachtung von Interaktionen, Modellbildung anhand theoretischer Vorkenntnisse etc.). Die Formanalyse hat sich, insbesondere auch in erstmaliger Kombination mit anderen Verfahren, in den beschriebenen Beispielen als geeignet erwiesen.

#### 10.4.2 Vorhersage

Bei der Vorhersage sind in der Regel, Daten zu verwenden, die für die Modellherstellung noch nicht benutzt wurden. Meist wird die Vorhersage durch Datensplits modelliert, wie ich dies beispielhaft anhand von Sensordaten durchgeführt habe [21]. Bei diesem Vorgehen stellen sich zwei Fragen, zum einen nach der Häufigkeit der Datensplits und zum anderen nach dem dann zu verwendenden Modell.

Gerade in dem Bereich der Medizin wäre die Prognosefähigkeit des hier verwendeten Modells bei dem nächsten Jahrgang direkt zu prüfen anstatt Datensplits zu verwenden. Bei den olfaktorischen Messungen stellt sich die Frage der Vergleichbarkeit von realer Vorhersage und der anhand der Daten simulierten (nur Messungen morgens, nur zu einer Jahreszeit etc.). Bei den Serienmördern ist es besonders schwer zur Vorhersage zu gelangen, da gerade ihr Verhalten als doch in der Regel meist gerade durch die Unkontrolliertheit und Unvorhersehbarkeit geprägt erscheint.

Davon ausgehend, dass die Formanalyse erwartungstreue Parameter der Verteilung liefert, können die getroffenen Aussagen gerade in der Medizin die bisherige radiologische Diagnostik um ein nachvollziehbares mathematisches Verfahren erweitern. Ein weiterer Ansatz in der Medizin wurde bereits im Abschnitt der dynamischen Formanalyse vorgestellt, in der die Formveränderung des Herzens für die Diagnostik verwendet werden kann. Weitere Anwendungen sind darüber hinaus in der Gerichtsmedizin denkbar<sup>1</sup>.

Bei den Serienmördern ist die zeitliche Veränderung und die geringe Datenmenge bei der zukünftigen Verwendung zu beachten. Die Übertragbarkeit auf Einzelfälle erscheint aufgrund der Besonderheiten der Täter (psychische Abnormalität etc.) äußerst fragwürdig. Zumindest die über die Jahre gewachsene Mobilität kann genauso wie die individuell sehr unterschiedliche Fähigkeit, sich im Raum zu bewegen, über die Normierung herausgerechnet werden. Nicht herausgerechnet werden kann bei dem jetzigen Vorgehen räumliche Besonderheiten und Begrenzungen wie die Inlandsituation von Berlin vor dem Mauerfall.

Die Auswertung der elektronischen Nasen ist von der verwendeten Nase abhängig.

---

<sup>1</sup> vgl. Rißquetschwunden auf S.30 in [84]

---

Selbst bei einem Hersteller sind Abweichungen unter den elektronischen Nasen zu erwarten. Diese Abweichungen können nur bedingt durch die Formanalyse abgefangen werden. Die Belastung mit Geruchssubstanzen über einen längeren Zeitraum führt zu sich unterschiedlich auswirkenden Alterungsprozessen auf die Sensoren. Die Formanalyse kann sich in diesem Bereich lediglich als ein Auswertungsverfahren etablieren, ihre Resultate in der vorliegenden Arbeit sind weitergehend kritisch zu prüfen.

### 10.4.3 Ausblick

#### Psychologie

In der Psychologie geht es um die Bildung von Profilen. Die Profile ergeben sich aus dem Verhältnis der einzelnen Itemantworten zueinander. So könnte anstatt der üblichen Scorebildung die Items zweidimensional abgetragen und mittels der Formanalyse bearbeitet werden. Es wird dann ein Profil derer gebildet, die erneut straffällig geworden sind, und derer, die nicht erneut strafrechtlich in Erscheinung getreten sind und miteinander verglichen.

Kritisch bei diesem Vorgehen ist die Anordnung der Items in der zweidimensionalen Ebene. Die Anordnung kann sich durch die Anordnung der Items an sich ergeben oder bereits theoretisch begründet werden.

#### Bevölkerungsstatistik

Beim demographischen Wandel, wie bereits von uns kritisch in [39] betrachte, stellt sich die Frage, wie anhand weniger Kennziffern eine Aussage zu treffen ist. So wird die Frage behandelt, ob aus den bereits vergangenen 14 Jahren eine Aussage zu der zukünftigen Entwicklung der Kriminalität zu rechtfertigen ist. Angesichts der niedrigen Fallzahl sind exakte Aussagen zur Entwicklung nicht zu erwarten. Für den Anwender würde es bereits reichen, festzustellen zu können, ob bei dem kommenden Jahr von einer steigenden oder sinkenden Rate der Kriminalität auszugehen ist. Auch in diesem Fall können unterschiedliche demographische Faktoren zweidimensional abgetragen werden und entsprechend ausgewertet werden.

#### Gesichtserkennung

Im Gegensatz zu den Neuronalen Netzen ist bei der Formanalyse die Verwendung der Landmarken nachvollziehbar. Von vornherein können theoretische Annahmen über die relevanten Landmarken in das Verfahren eingefügt werden. Eine mögliche Kombination Neuronaler Netze nach eingehender Verarbeitung durch die Formanalyse könnte zu einer entsprechenden Verbesserung in der Gesichtserkennung führen.

---

## **Ökonomie und Finanzmathematik**

Die mittlere zu erwartende Funktion bietet einen neueren Ansatz zur Bewertung eines Portfolios. Die hier aufgezeigte Kombination von Neuronalen Netzen mit der Stochastik führen zu einer neuen Sichtweise auf die stochastischen Prozesse und zu einer Interpretierbarkeit der bisher undurchsichtigen neuronalen Netze [46]. Dieser Ansatz geht über die Beschäftigung mit der Form hinaus und bietet eine Brücke zwischen Numerik und Stochastik.

## **Weitere Anwendungen**

Von mir wird in den folgenden Jahren vor allem die dynamische Formanalyse weiter zu entwickeln sein, die ein erhebliches Potential in allen Bereichen aufweist und dem Anwender die Möglichkeit bietet, die Veränderungen nicht nur anhand von Kennziffern darzustellen, sondern auch optisch zu erkennen.

Im Weiteren wird die Kombination unterschiedlicher mathematischer Verfahren zur Problemlösung weiter untersucht werden, die bereits in dieser Arbeit und in vielen anderen Bereichen erfolgsversprechend erscheint.

---

# Literaturverzeichnis

- [1] 3D-Doctor, <http://www.ablesw.com/3d-doctor/>
- [2] Agresti, Alan, 1990, Categorical Data Analysis, Wiley, New York 1990
- [3] Backhaus, K., Erichson, B., Plinke, W., Weiber, R., 2000, Multivariate Analysemethoden, Springer, Berlin, Heidelberg, New York
- [4] Innovations-Report, 2010,  
[http://www.innovations-report.de/html/berichte/biowissenschaften\\_chemie/bericht-35511.html](http://www.innovations-report.de/html/berichte/biowissenschaften_chemie/bericht-35511.html)
- [5] Bishop, C.M., 1995, Neural networks for pattern recognition, Clarendon Press, Oxford.
- [6] BKA (Hrsg.),  
<http://www.bka.de/kriminalwissenschaften/fotofahndung/pdf/fotofahndung-abschlussbericht.pdf>
- [7] Bookstein, F.L., 1986, Size and shape spaces for landmark data in two dimensions, Statistic Sciences 1; p.181-242.
- [8] Bookstein, F.L., 1996, Biometrics, biomathematics and the morphometric synthesis, Bulletin of Mathematical Biology 58; p.313-365.
- [9] Breslow, N., Olshan, A., Beckwith, J.B., Green, D.M., 1993, Epidemiology of Wilms tumor, Med. Pediatr. Oncol. 21, S. 233-236
- [10] Cao, F., Morel, J.-M., Sur, F., Lisani, J.-L., Musé, P., 2008, A Theory of Shape Identification, Springer, Berlin, Heidelberg
- [11] Campbell, Duncan, Are terrorists using hidden messages?, 31.10.2001,  
<http://www.heise.de/tp/r4/artikel/11/11005/1.html>
- [12] Coppes, MJ, Campbell CE, Williams, BRG, 1995, Wilms Tumor: Clinical and Molecular Characterization, Austin Texas USA: RG Landes Company, S.1-55
- [13] Dern, H., Frönd, R., Straub, U., Vick, J., Witt, R., Geografisches Verhalten fremder Täter bei sexuellen Gewaltdelikten, BKA (Hrsg.), Wiesbaden 2004

- 
- [14] DIN EN 13725, 2003-07
  - [15] Uni Düsseldorf (Hrsg.), <http://www.uni-duesseldorf.de/AWMF/II/025-004.html>, 15.09.2006
  - [16] Ebert, A.; Gravou, C., Stümpfl, M., Rösch, W., 2003, Das Nierenzellkarzinom im Kindesalter, S.263-268 in "Der Urologe A" Bd. 42 Nummer 2, Springer-Verlag, Berlin, Heidelberg
  - [17] Favoro, Paolo, and Soatto, Stefano, 2007, 3D-Shape Estimation and Image Restoration, Springer, Berlin.
  - [18] Faller, Adolf, 1988, Der Körper des Menschen-Einführung in Bau und Funktion, Thieme, Stuttgart, New York
  - [19] R.A.Fisher, 1936, The statistical utilization of multiple measurement', Annals of Eugenics, Bd. 8, 1936, S. 376-386
  - [20] Georgii, H.-O., 2002, Stochastik, Gruyter, Berlin, New York
  - [21] Giebel, S.M., 2007, Schätzung der Geruchsstoffkonzentration anhand von Sensorarrays: Vergleich der Logistischen Regression mit den Neuronalen Netzen, VDI Tagungsband "'Gerüche in der Umwelt", Bad Kissingen 2007
  - [22] Giebel, S.M., 2006, Vorhersage und Erklärung von Rückfälligkeit jugendlicher Straftäter, Bulletin de la Société des Sciences Médicales du Grand-Duché de Luxembourg, Numéro Special 2006, S.233-247
  - [23] Giebel, S.M., Vortrag,  
<http://www3.iam.metu.edu.tr/juergenlehn/giebel.pdf>
  - [24] Giebel, S.M., 2008, Relapse of juvenile offenders between late 1996 and early 2000, Nuove Esperienze di Giustizia Minorile, 2-2008, S.99-109
  - [25] Giebel, S.M., 2008, Evaluation der Suchttherapie im. Maßregelvollzug Hadamar, Research in Forensic Psychiatry, Vortrag 19.-21.06.2008
  - [26] Giebel, S.M., 2008, Statistische Analyse der Form von Nierentumoren bei Kleinkindern, Bulletin de la Société des Sciences Médicales du Grand-Duché de Luxembourg, Numéro Special 1/08
  - [27] Giebel, S.M., Frechen, F.-B., Franke, W., 2008, Anwendung der "Formanalyse" auf Geruchsprofile, DepoTech2008, Lorber K.E. (Hrsg.), VGE Verlag, Essen
  - [28] Frechen, F.B., Franke, W., Giebel, S., 2009, Biofiltration: From past to future - scientific view- Biotechniques for Air Pollution Control, available on <http://hrz-vm162.hrz.uni-kassel.de/web/SiwawiDokumente/pdf/publikationen>.

- 
- [29] Frechen, F.B., Giebel, S., Anwendung der logistischen Regression auf Geruchssensoren zur Bestimmung olfaktometrischer Messungen, 3rd Statistical Days Luxembourg, available on <http://sma.uni.lu/stat3/files/Frechen.pdf>
  - [30] Frechen, F.-B., The new VDI Guideline 3885/1 for OEC measurement and why we might need electronic noses, Nose2010, Florenz, 22-24.September 2010
  - [31] Giebel, S., 2007, Analyse of the spatio-temporal patterns of serial murderers (Analyse des räumlichen Verhaltens von Serienmördern), 3rd Statistical days, available on <http://sma.uni.lu/stat3/files/Giebel2.pdf>.
  - [32] Giebel, S., 2007, Statistical Analysis of the shape of renal tumors in childhood, Diploma thesis, University Kassel.
  - [33] Giebel, S.M, Schiltz, J., Schenk,J.-P., 2009, Application of Shape Analysis on renal tumors, Vilnius, 30.Juni-3.Juli 2009
  - [34] Giebel, S.M, Schiltz, J., Schenk,J.-P., 2009, Application of Shape Analysis on renal tumors, Porto Heli, ICCRA 3, 16.-18.Juli 2009
  - [35] Giebel, S., Schiltz, J., and Schenk, J.P., 2009,Application of Statistical Shape Analysis to the Classification of three dimensional renal tumors appearing in early childhood; available on <http://www3.iam.metu.edu.tr/juergenlehn/giebel.pdf>
  - [36] Giebel, S.M., Rainer, M., 2008, Forecast of stock prices by using neural networks in comparison to log-normal estimation, <http://sma.uni.lu/stat4/files/Rainer.pdf>
  - [37] Giebel, S., Schiltz, J., Schenk, J.-P. ,2010, Application of Shape Analysis on renal tumors in 3D, IEEE, Hibit Conference, Antalya
  - [38] Giebel, S.M., Schiltz, J., Graf, N., Schenk, J.-P., 2010, Formanalyse in der Magnetoresonanztomografie-Landmarken und Objektdifferenzierung bei retroperitonealen Tumoren im Kindesalter, Bulletin de la Société de Sciences Médicales du Grand-Duché de Luxembourg, Numéro Special 1/10
  - [39] Giebel,S.M., Kemme, S., 2010 Weniger Jugendliche, weniger (Jugend-)kriminalität, weniger Polizei(-bedarf)? In Frevel, B., Bredthauer, B. (Hrsg.): Empirische Polizeiforschung XII: Demografischer Wandel und Polizei. Frankfurt a.M.: Verlag für die Polizeiwissenschaft: Schriften zur Empirischen Polizeiforschung, S. 71-97
  - [40] Giebel, S.M., Rainer, M., 2010, Stochastic processes adapted by neural networks exhibiting possible transitions to chaos with application to climate, energy, and finance, 10th Iranian Statistical Conference, Tabriz

- 
- [41] Giebel, S.M., Hosser, D., 2010, Criminal careers of young adult offenders after release, Operational Research, München, 2.September 2010
  - [42] Giebel, S.M, Rainer, M., 2010, Stochastic processes adapted by neural networks with application to climate, energy, and finance, Special volume in Applied Mathematics and Computation (in Druck)
  - [43] Giebel, S.M., Rainer, M., 2010, Climate models and commodity pricing with stochastic processes adapted by neural network, PCI 2010 - 3rd. Int. Conf. Problems of Cybernetics and Informatics, Baku 6-8.9.2010
  - [44] Giebel S., Schiltz J., Nourkami N., Leuschner I., Graf N., Schenk J.-P., 2010, 3D-Formanalyse von retroperitonealen Tumoren im Kindesalter: Entwicklung einer statistischen 2D/3D-Methodik und Anwendung dieser mit MRT-Datensätzen., 47. Internationale Jahrestagung der Gesellschaft für Pädiatrische Radiologie
  - [45] Giebel, S.M., Schiltz, J., Schenk J.-P., 2010, Differentiation Tests for Three Dimensional Shape Analysis, [http://www.smta.net/images/Book\\_of\\_abstracts\\_smta2010\\_final\\_1.pdf](http://www.smta.net/images/Book_of_abstracts_smta2010_final_1.pdf)
  - [46] Giebel, S., Rainer, M., 2010, Stochastic estimation of energy resources (wind) and gas prices, Energy & Finance Conference, 2nd International Ruhr Energy conference, Essen, 6.-8.Okt
  - [47] Giebel, S.M, 2010, Zur Anwendung Neuronaler Netze in den Sozialwissenschaften, Verlag Dr. Kovac, Hamburg
  - [48] Giebel, S., 2010, Evaluation des saarländischen Jugendstrafvollzugs, Forschungsantrag im Auftrage des Justizministeriums
  - [49] Giebel, S., 2010, Evaluation des Maßregelvollzugs Hadamar, Dissertation an der Universität Kassel
  - [50] Harbort, S., 2008, Begegnung mit dem Serienmörder, Droste, Düsseldorf
  - [51] Harbort, S., Giebel, S., 2010, Daten zu Raubmördern in Form von Kartenmaterial, unveröffentlicht
  - [52] Heidegger, M., 1993, Sein und Zeit. Tübingen: Max Niemeyer Verlag
  - [53] IFI, 2007, <http://www.ifi.unizh.ch/>
  - [54] Kinderkrebsinfo, 2006, [www.kinderkrebsinfo.de](http://www.kinderkrebsinfo.de)
  - [55] McCulloch, W.S., Pitts, W., 1943, A logical calculus of the idea immanent in neural nets, Bulletin of Mathematical Biophysics, S. 115-137.
  - [56] Metzler, W., 1998, Nichtlineare Dynamik und Chaos, Teubner Sachbücher, Stuttgart, Leipzig



- 
- [57] Moosbrugger, H., Frank, D., 1992, Clusteranalytische Methoden in der Persönlichkeitsforschung, Verlag Hans Huber, u.a. Bern, S.16
  - [58] Najarian, Kayvan, On stochastic stability of dynamic neural models in presence of noise, Proceedings of the 4th International Conference on dynamical system and differential equations, May 24 -27, 2002, Wilmington, NC, USA, S. 656-663
  - [59] Coppes, M.J., Campbell C.E., Williams, B.R.G., 1995, Wilms Tumor: Clinical and Molecular Characterization, Austin Texas USA: RG Landes Company, p.1-55
  - [60] Dryden, I. L; Mardia, K. V., 1998, Statistical Shape Analysis. Wiley, Chichester
  - [61] Furtwängler, R., Schenk, J.P., Graf, N. u.a., 2005, Nephroblastom- Wilms-Tumor, Onkologie 11.
  - [62] Fréchet, M., 1948, Random elements in the metric space (Les éléments aléatoires de nature quelconque dans un espace distancié), Annales de l'Institut Henri Poincaré, Vol. 10, pp.215-310
  - [63] Reichenbach, H, Gesammelte Werke, Bd.1/ Bd. 4/ Bd. 7, Vieweg Verlag, Braunschweig 1977
  - [64] Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986, Learning internal representation by error propagation, in Rumelhart, D.E.; McClelland, J.L.; PDP Research Group (Hrsg.), Parallel. Distributed Processing, MIT Press, Cambridge, S.318-36
  - [65] Peirce, C.S., <http://www.helsinki.fi/science/commens/peircetexts.html>
  - [66] Modern English textbooks and translations prefer "theory of Form" to "theory of Ideas", but the latter has a long and respected tradition starting with Cicero and continuing in German philosophy until today, and some English philosophers prefer this in English too. See W D Ross, Plato's Theory of Ideas (1951) and this reference site.
  - [67] Graf, N., 2003, Urologe A 43:421
  - [68] Lienert, G.A. ,1969, Die Konfigurationsfrequenzanalyse als Klassifikationsmittel in der klinischen Psychologie, In: Irle, M. (Hrsg.): Bericht über den 26. Kongreß der Deutschen Gesellschaft für Psychologie 1968 in Tübingen, Göttingen 1969
  - [69] Dryden, I.L., Mardia, K.V., 1998, Statistical Shape Analysis, Wiley, Chichester

- 
- [70] Mattheou, K., Karagrigou, A., Neophytou, P., 2009, Multivariate Statistical Analysis for the incidence of HPV-infection in Cypriot male population, Kitsos, C., Alexandros, G.R., Biebler, K. E. (Hrsg.), Cancer Risk Assessment, Selected papers from ICCRA3, Shaker Verlag, Aachen
  - [71] Kendall, D.G., 1977, The diffusion of shape, Adv. Appl. Probab. 9, 1977
  - [72] Kendall, D.G., 1984, Shape manifolds, Procrustean metrics and complex projective spaces, Bulletin of the London Mathematical Society, 16, S.81-121.
  - [73] Kendall D.G., Kendall W.S., 1980, Alignment in two dimensional random sets of points, Advances in Applied probability 12, S.380-424.
  - [74] Kohonen, T., 1984, Self- Organization and Associative Memory, Springer, u.a. Berlin
  - [75] Krauth, J., 1993, Einführung in die Konfigurationsfrequenzanalyse, Beltz, Weinheim
  - [76] Mardia, K.V., 1977, Mahalanobis distance and angles. In: Krishnaiah, P.R. (ed) Multivariate Analysis IV, p.495-511, Amsterdam: North Holland.
  - [77] Minsky. M.; Papert, S., 1969, Perceptrons, an introduction to computational geometry, MIT Press, Cambridge, MA
  - [78] Österreichische Gesellschaft für Chirurgische Onkologie, [http://www.aco-asso.at/manual/aktuell/kinder/3/03\\_07.html](http://www.aco-asso.at/manual/aktuell/kinder/3/03_07.html)
  - [79] Sachs, L., 1991, Angewandte Statistik, Springer Verlag Berlin, Heidelberg, New York, S.129
  - [80] Schenk, J.P, Fortschr. Röntgenstr. 178:38
  - [81] Schenk, J.P, 2008, Eur. Radiol. 18:683, 2008
  - [82] Schwind, H.D., 2008, Kriminologie, Kriminalistik Verlag, Heidelberg, S.310-349
  - [83] Small, C. G., 1996 The Statistical Theory of Shape. Springer-Verlag, New York.
  - [84] Spann, M., 1987, Angewandte Gerichtsmedizin, Urban&Schwarzenberg, Wien, München, Baltimore
  - [85] SPSS, <http://webs.edinboro.edu/EDocs/SPSS/SPSS%20Classification%20Trees%202013.0.pdf>
  - [86] Stoyan, D., On means, medians and variances on random compact sets, Geobild, Berlin 1989
  - [87] UICC, 2006, <http://www.uicc.org/index.php?id=1167>

- 
- [88] Wilms, M., 1889, Renal tumors (Die Mischgeschwülste der Niere), Verlag von Arthur Georgi, Leipzig, p.1-90.
- [89] Ziezold, H., 1974, On expected figures and a strong law of large numbers for random elements in quasi- metric spaces, Trans. 7<sup>th</sup> Prague Conference Inf. Th., Statist. Dec. Funct., Random Processes, Vol. A. Reidel, Dordrecht, Prag, S. 591 - 602
- [90] Ziezold, H., 1994, Mean Figures and Mean Shapes Applied to Biological Figure and Shape Distributions in the Plane, Biometrical Journal, 36, p. 491-510.
- [91] Ziezold, H., 2003, Independence of Landmarks of Shapes. Mathematische Schriften Kassel, Heft 03